

Market Discipline in Banking Regulation: Theory and Evidence from Switzerland

Matteo Facchinetti

Università della Svizzera Italiana and Swiss National Bank

Dissertation submitted to the

FACULTY OF ECONOMICS

of the

UNIVERSITÀ DELLA SVIZZERA ITALIANA

under the supervision of

Prof. Dr. Giovanni Barone-Adesi

Fall 2007

Preface

The opinions expressed in this thesis are those of the researcher and do not necessarily reflect those of the Swiss National Bank or its staff. I am grateful to my supervisor Giovanni Barone-Adesi for his advice and the helpful discussions. I am also indebted to my co-supervisor Enrico De Giorgi and to Andrea Sironi for their valuable comments. Special thanks go to Urs W. Birchler, Jürg Blum, and my other colleagues at the Swiss National Bank. Finally, I would like to thank all my family for their endless support.

Contents

Preface	iii
1 Executive summary	1
1.1 Introduction and motivation	1
1.2 Papers	2
1.3 Conclusions	7
2 The prospects of market discipline for banks. Evidence from the Swiss bond market	9
2.1 The theoretical model	12
2.1.1 One-period game	14
2.1.2 Infinitely repeated game	16
2.1.3 Discussion	18
2.2 Empirical analysis	20
2.2.1 Variable description	21
2.2.2 The empirical model	28
2.2.3 Data set and sample characteristics	31
2.3 Results	32
2.3.1 Linear specification	32
2.3.2 Discussion and robustness tests	36
2.3.3 Nonlinearities	41
2.4 Conclusions	43
2.5 Appendix	46
2.5.1 Conditions for the acceptance of the debt contracts	46
2.5.2 Expected value of equity in the infinitely repeated game	49
2.5.3 Introducing concavity into the manager's objective function	51
2.5.4 Variable definition	53
2.5.5 The variance covariance matrix	55
2.5.6 Complete estimation results	55
3 Is direct market discipline in banking really effective and necessary? A comparison between public and private banks	57
3.1 Related literature	60

Contents

3.2	A simple static model	61
3.2.1	Market discipline	62
3.2.2	No market discipline	67
3.3	Some remarks in a dynamic framework	68
3.4	Empirical analysis	71
3.4.1	Methodology and objectives	71
3.4.2	Model specification	72
3.4.3	Data set	77
3.5	Results	79
3.5.1	Cobb-Douglas model	80
3.5.2	Procyclicality	83
3.5.3	Robustness tests	84
3.5.4	Discussion	86
3.6	Conclusions	89
3.7	Appendix	89
3.7.1	Complete results of the Cobb-Douglas model	89
3.7.2	Risk-taking incentives and debt	91
3.7.3	Effect of the cycle on expected revenues	91
4	The endogeneity pitfall in using market signals as triggers for prompt corrective action	93
4.1	The informational content of prices for bank debt and equity	95
4.2	The double endogeneity problem	96
4.2.1	A potential paradox	96
4.2.2	Fragile equilibrium	96
4.2.3	Previous literature	97
4.3	The model	98
4.3.1	Assumptions	98
4.3.2	Optimal strategies	101
4.3.3	Equilibrium: Benchmark cases	104
4.3.4	Equilibrium: Asymmetric information and competitive insiders	104
4.3.5	Equilibrium: Asymmetric information and monopolistic insiders	113
4.4	Discussion	116
4.4.1	The mechanics behind the endogeneity problem	116
4.4.2	A way around endogeneity: Price neutral intervention	117
4.5	Conclusions	119
4.6	Appendix	120
4.6.1	The trigger value of π_0	120
4.6.2	Equilibrium price with monopolistic insiders	121

4.6.3 Equilibrium with price neutral intervention	122
Bibliography	123

List of Figures

2.1	Average subordinated bond spreads	41
2.2	Objective function of the manager in the static model	52
2.3	Objective function of the manager in the dynamic model	52
3.1	Market share in the Swiss credit market	88
4.1	Time line	101
4.2	Equilibrium without quasi-fixed cost ($K = 0$).	106
4.3	Reaction function of the authority with quasi-fixed cost	107
4.4	Partially revealing equilibrium with quasi-fixed cost.	110
4.5	Fully revealing equilibrium with quasi-fixed cost.	111

List of Tables

2.1	Summary statistics	31
2.2	Base-model estimation results	34
2.3	Panel sensitivity	38
2.4	Unobserved heterogeneity	39
2.5	Base-model with no leads	40
2.6	Importance of nonlinearities	42
2.7	Spline model with general knots	44
2.8	Spline model with bank-specific knots	45
2.9	Variable description	54
2.10	Complete base-model estimation results	56
3.1	Return distribution of the portfolios	61
3.2	Empirical implications of the static model	68
3.3	Description of the explanatory variables	73
3.4	Summary statistics	78
3.5	The Cobb-Douglas model	81
3.6	Procyclicality	84
3.7	Additional measures of risk and performance	86
3.8	Complete results of the Cobb-Douglas model	90

1 Executive summary

1.1 Introduction and motivation

Bank supervisors never had an easy job: Banks and in particular the largest bank holding companies are extremely complex and opaque institutions. In recent years, the proliferation of financial innovations as well as the increasing globalization and concentration in the financial sector have made bank supervision even more challenging. Timely identification of problems at individual banks and of disequilibria building up in the whole financial system turns out to be a daunting task.

Markets represent a key factor affecting the development of the banking industry. From the perspective of bank supervisors, however, markets play an ambiguous role. On the one hand, they are the driving force behind the increasing complexity of the banking industry. Financial markets, in particular, offer today an extraordinary choice of sophisticated instruments to diversify, hedge, but also to conceal risks. On the other hand, markets may help supervisors in promoting a safe and sound financial system through market discipline. Market discipline can take two forms. Economists speak about *direct* market discipline when banks that are prone to assume excessive risks find it difficult or expensive to refinance those risks. Markets exert *indirect* discipline when the terms at which banks refinance reveal valuable information on investors' perception of bank quality to supervisors. The Basel Committee on Banking Supervision, uniting central bankers and supervisors from industrialized countries, recognizes the potential of bank-specific market information. Its revised minimum international capital framework for banks (commonly known as "Basel II") relies on market discipline as the third of three pillars. The set of disclosure requirements for banks specified in "Basel II" is thought to bolster up direct market discipline by investors. At the same time, the improved transparency should allow a more risk adequate pricing of bank financial securities and hence provide better market information for the national supervisory review process under the second of the three pillars.

In Switzerland, few attempts have been made to assess the role and potential of market discipline in banking supervision. The main aim of my thesis is to fill this gap. In the first two papers, I explore to what degree the market disciplines Swiss banks, both directly and indirectly. The third paper examines some implications of the use of market information by the supervisor, in particular in the context of prompt corrective actions.

1.2 Papers

I) The prospects of market discipline for banks. Evidence from the Swiss bond market

Market prices are expected to fulfil a key economic function in that they process and aggregate investors' information. According to the (semi-strong) efficient market hypothesis, prices reflect all public available information. In particular, market data such as observed yields on bank bonds or CDS premia are supposed to convey signals about bank quality and risk. As this is a first necessary condition for the existence of market discipline, both direct and indirect, in my first paper I analyze to what extent this market monitoring hypothesis is consistent with empirical evidence in Switzerland. In addition, I compare the information contained in financial prices with that available to the Swiss supervisor.

Many economists already provided empirical evidence that investors are risk-sensitive and require an adequate risk premium for the money they invest or lend to banks. However, this research mainly focused on large banks in the US (see Flannery, 1998 for an overview of the early literature and DeYoung, Flannery, Lang, and Sorescu, 2001; Swidler and Wilcox, 2002; Fan, Haubrich, Ritchken, and Thomson, 2003; Krainer and Lopez, 2004; Krishnan, Ritchken, and Thomson, 2006; Curry, Elmer, and Fissel, 2007 for more recent research), in Europe (Persson, 2002; Sironi, 2002, 2003; Gropp, Vesala, and Vulpes, 2006) and in Japan (Brewer, Genay, Hunter, and Kaufman, 2003; Imai, 2007). Some researchers (for instance Berger, Davies, and Flannery, 2000; Evanoff and Wall, 2001; Gunther, Levonian, and Moore, 2001; Cannata and Quagliariello, 2005) go a step further and suggest that internal supervisory information and market information are complementary sources of supervisory intelligence. In other words, even banks' overseers may learn something by looking at market data.

Several features differentiate my study from the existing literature.

First, I analyze credit spreads of bonds exclusively issued by Swiss banks. In Switzerland the use of market data as information source seems particularly tempting, since banking supervision is an indirect or two-stage process. With the exception of the two big banks UBS and Credit Suisse Group, the Swiss supervisor, the Swiss Federal Banking Commission (SFBC), does not base its monitoring process on on-site supervision but on external auditors' reports. Hence, compared to supervisory authorities in other countries such as the US, where on-site examinations are more frequent, the SFBC may have less direct information regarding bank solvency and risk. To my knowledge, empirical work on market discipline in Switzerland is limited to Birchler and Maechler (2002), who look at deposit-based funding and find that uninsured depositors tend to withdraw money from banks that experience a deterioration of their risk profile. In particular, nothing has been done on the role of the capital market, which, because of the sophistication of its participants, may be particularly important in this context.

Second, in the aforementioned literature only large banks are usually taken into account. By the nature of the Swiss market, I also consider other bank groups, in particular relatively small regional

banks and state-owned institutions. This is an interesting issue as the existence of market discipline at smaller institutions may facilitate the allocation of the usually limited resources available to the supervisor. The other side of the coin is that one has also to deal with problems such as poor asset liquidity and data quality, since few small Swiss banks issue financial securities on a frequent basis.

The explicit use of a model that tries to describe the relationship between risk and bond prices is another particular feature of this paper. Its main feature is that it evaluates the interactions between all the parties concerned with the risk profile of a bank, namely the shareholders/managers and various types of debt-holders. In particular, I endogenize the risk decision process of the bank managers. Since other researchers considered only the relationship between the risk profile of a bank and bond spreads leaving the risk choice unexplained, my paradigm gives some additional insights when interpreting the estimation results.

Consistent with international evidence, my findings suggest that market signals are informative. Debenture yields react to changes in the risk profile of the issuer. Bonds issued by cantonal banks – public banks whose liabilities are guaranteed by the state – represent an exception, though. The existence of a state guarantee almost destroys the informational content of financial market data. The natural explanation for this result is that investors feel they are not at risk.

Even though market data contain some information on the risk profile of banks, I find little empirical evidence that this information is useful for the Swiss supervisor. In other words, it seems that the market does not know more or learn something about a particular institution significantly before the SFBC. However, this conclusion rests on the strong assumption that market prices do not play any role in the supervisory process for the period under scrutiny.

II) Is direct market discipline in banking really effective and necessary? A comparison between public and private banks

Given that market prices generally react to changes in the risk profile of a bank, an additional question emerges with respect to direct market discipline. Following the distinction between market monitoring and market influence emphasized by Bliss and Flannery (2002), in my second paper I examine whether the risk-sensitive supply of funds has an effect on bank behavior, i.e., whether the market can directly discipline banks by reducing managers' incentives for assuming excessive risks. In order to test this hypothesis, I compare the lending policy and performance of cantonal banks, which should not be subject to market discipline as they enjoy a state guarantee, with the lending policy and performance of private banks. As illustrated by Merton (1977) in the context of deposit insurance, safety nets have moral hazard consequences in that they increase the risk appetite of banks above efficient levels, also in the absence of bankruptcy costs. Thus, in the presence of direct market discipline, moral hazard combined with superficial monitoring by the main equity-holder (the state) may induce the managers of public banks to be more risk-seeking and reckless in lending money than managers of private banks.

Several researchers compared the performance of public banks with the performance of other bank groups in the past. Rime and Stroh (2003) examine the efficiency of Swiss banks between 1996 and

1 Executive summary

1999, the main aim being to explain the consolidation process in the Swiss banking industry. Among other things they reject the hypothesis that cantonal banks are more profit-efficient than other Swiss bank groups that are privately owned. By focusing on the German and Austrian banking sectors Altunbas, Evans, and Molyneux (2001) and Hauner (2005) come to similar conclusions. Although Hauner finds that public banks are more efficient, he attributes the result to the cheaper funding costs public institutions enjoy thanks to the state guarantee. Bichsel (2006) analyzes the role of cantonal banks from the perspective of competition and concludes that they do not deviate from the typical conduct in the Swiss banking sector. By looking at interest rates on deposits, Bichsel also provides empirical evidence that cantonal banks have lower-than-average funding costs. He suggests that this discount may be explained by the state guarantee. Birchler and Maechler (2002) find that depositors of cantonal banks are less responsive to changes in bank fundamentals than those of Swiss private banks. Sironi (2002, 2003) and Pop (2006) arrive at similar conclusions in the area of capital market-based funding. Sironi (2003) studies subordinated debt issues in Europe and observes that, after controlling for the bank's risk, bonds issued by public banks trade at an average discount premium of about 40 basis points. Additionally, Sironi's findings corroborate the hypothesis that debt investors are able to discriminate between the different risk profiles of private banks, but they care less about risk when pricing bonds issued by public or guaranteed institutions. Sironi (2002) and Pop argue that because of the state guarantee the average European and Japanese bank bond trades at a lower spread than its North American counterpart. Finally, Flannery and Sorescu (1996) show that US bank investors became risk-sensitive only when the regulator made clear that they were not protected in the case of default. This line of research combined with the evidence provided by my first thesis paper suggests that in the presence of state guarantees investors hardly price idiosyncratic bank risk and therefore a prerequisite of direct market discipline is not satisfied. However, there is almost no study that goes one step further and examines whether this risk-insensitive supply of funds induces public banks to be less cautious in lending money than other bank groups. One major exception is the paper by Iannotta, Nocera, and Sironi (2007). The authors compare risk and performance of European public, mutual and private banks and find that public banks are, on average, less profitable but at the same time they tend to have lower quality assets. In Switzerland this issue is of particular relevance since cantonal banks constitute the second biggest group of banks. Furthermore, given the evidence that cantonal banks have lower financing costs thanks to the state guarantee, the fact that they are not more profit-efficient than other Swiss banks is left unexplained.

Another branch of the literature examines the empirical relevance of direct market discipline, especially in the US. Those studies typically focus on the relationship between bank risk and (subordinated) debt contracts. The evidence from this literature is mixed. Bliss and Flannery (2002) analyze whether bank holding companies' stock and bond prices affect management actions but their results provide little support for investor influence. Goyal (2005), on the other hand, observes that banks with greater risk-taking incentives include more restrictive covenants in their debt contracts, which should reduce yields spreads but also managerial flexibility to increase risk. Nier and

Baumann (2006) find a positive relationship between bank equity capital and several proxies for the strength of market discipline (depositor protection, uninsured funding and transparency) in a panel of banks from different countries. Yet, much less work has been done to analyze the effects of market discipline directly on the asset side of the balance sheet and on lending policy in particular.

The second paper of my thesis tries to close all these gaps and link these two branches of the literature. In other words, I study the effectiveness of market discipline on the asset side of the balance sheet by assessing the relative efficiency of the lending policy of public banks in comparison with that of private banks. One problem that arises when one compares private and public banks from the perspective of market discipline is that moral hazard distorts the incentives only when the managers act on behalf of equity-holders and maximize shareholder value and the benefits that arise from limited liability. However, in the absence of distortions, public banks are expected to pursue objectives that are more welfare oriented, for instance to sustain the local economy. Hence, I develop a simple theoretical model that allows to analyze the impact of market discipline and (public) bank's objectives on lending performance. In particular, I show that when market discipline is effective, private banks never assume more risks than public banks, independently from the objective function of the latter. Private banks are likely to be less risky and more profitable, especially when the government does not exercise sufficient control over the management of public banks.

In a second step, the implications derived from the model are tested on Swiss bank data. Thanks to its particular structure – cantonal banks constitute Switzerland's second biggest group of banks – the Swiss banking sector offers an ideal background for comparing public and private banks from the perspective of direct market discipline. The empirical analysis provides strong evidence that cantonal banks, on average, have to build up more provisions and write down more loans than comparable private banks. In other words, it seems that public banks assume more credit risks. Accordingly, gross interest revenues tend to be higher at public banks. After controlling for increased credit risk, however, regional banks usually perform better than cantonal banks, even in absolute terms. Thus, I conclude that the market exercises some direct discipline on private banks and that public banks may lend inefficiently.

III) The endogeneity pitfall in using market signals as triggers for prompt corrective action

As market information in the form of data on the yields of bank equity and of subordinated debt is available for many banks in a number of countries, central banks and supervisory authorities are confronted with another question: Even if this data contains valuable information, to what extent – and how – should policy decisions in the area of banking and systemic stability be influenced by observed financial prices? One example on how supervisors could benefit from the information content of market prices in their assessment of banks' risk has been mentioned by the former chairman of the Federal Reserve Board, Alan Greenspan: "Significant changes in a banking organization's debt spreads, in absolute terms or compared with peer banks, can prompt more

1 *Executive summary*

intensive monitoring of the institution” (Greenspan, 2001). While supervisors in some countries do not rely on market information at all, some economists went to the other extreme by advocating subordinated debt spreads as automatic triggers for supervisory action (see, e.g., Calomiris, 1999 or Herring, 2004).

Up to now, the Basel Committee on Banking Supervision has made no explicit recommendation on the use of market data. In the absence of international principles and of practical experience, national authorities have little guidance on how to deal with bank-specific market data in their supervisory and system stability approaches. In the third essay with Urs W. Birchler as co-author, we highlight a basic informational problem that may arise when supervisors tie their actions to market prices. In particular, we show that such a policy may distort the informational content of financial prices and even lead to a market breakdown.

Existing research has already proved that prices are not always efficient aggregators of information a la Grossman (1976). In particular, several studies illustrate how information based trading may lead to a market breakdown or informationally inefficient prices (see, e.g., Hellwig, 1980; Milgrom and Stokey, 1982; Verrecchia, 1982). Following this branch of the literature, we use a dynamic rational expectations model to investigate the trading strategy of investors endowed with private information about the value of a risky asset. As opposed to the mentioned research, however, the distribution of the liquidation value of the risky asset is endogenous in our model. An additional agent, the authority, can affect this distribution if it is in its interest to do so after observing how trading evolves. As a result, additional market frictions are induced by moral hazard: The authority acting on the basis of market signals affects the fundamentals underlying the trading decisions of the investors and, consequently, the trading decisions themselves. Let us take for instance the credit spread of a bond issued by a bank. Intuitively, when investors think that the regulator will always bail out the bank should the spread exceed some threshold, they will anticipate the regulator’s reaction and price the bond accordingly. This may distort the incentives of the investors and limit the amount of information flowing in bond prices, even when investors are competitive. In particular, we show that problems emerge when the authority has quasi-fixed costs of intervention and intervention occurs at no cost for the investors. Finally, we illustrate how to overcome this endogeneity problem through an appropriate security design.

The mutual endogeneity of prices and price-based action was already noted by Krugman (1991) in the area of exchange rates. In the debate about whether central banks should react to asset price developments (to bubbles, in particular), Cecchetti, Genberg, and Wadhwani (2002) also note that the mere expectation of a central bank’s intent to react to asset prices would already change their dynamics. Somewhat surprisingly, in the context of bank supervision, this endogeneity effect was long overlooked. DeYoung, Flannery, Lang, and Sorescu (2001) observe that after an unexpectedly poor outcome of on-site examination bond spreads of financially troubled US banks tend to decrease, rather than to increase. They conclude that a poor result of an examination is more likely to trigger corrective action from the supervisor and that “the anticipated regulatory response frequently dominates the information’s implications about current bank conditions.”

(DeYoung, Flannery, Lang, and Sorescu, 2001). Shin (2002) notes that banks' asset values were affected by the endogenous response of authorities during a currency crisis. Bond, Goldstein, and Prescott (2006) show that because of the endogeneity problem the effectiveness of market based intervention strongly depends on the information gap between authority and the market. In their model, however, the action taken by the authority aims at reducing the agency costs between the managers of the bank and its equity and debt-holders. Moreover, the action set of the authority is binary, i.e., it can either intervene or not intervene. In our model the authority tries to minimize the probability of default of the bank and the degree of intervention is a continuous variable. Distortions caused by the endogeneity problem are emphasized also in Lehar, Seppi, and Strobl (2006). They show how the regulatory use of market data reduces investors' incentives to acquire costly information about the risk profile of the bank. At the same time such a policy may induce banks to increase risk taking. Yet, also in Lehar, Seppi, and Strobl the action set of the authority is binary and their model focuses on the interaction between policy response and information acquisition, rather than the efficient representation of existing information in prices. As opposed to the mentioned papers, however, we do not explicitly model the risk choice of the bank as well as the problem of the investor when collecting information is costly.

1.3 Conclusions

To sum up, my analysis provides evidence for the existence of some market monitoring and direct market discipline in the Swiss banking sector, at least for banks that do not benefit from a state guarantee. Accordingly, appropriate measures aimed at strengthening the effectiveness of direct market discipline – in particular through the third pillar of “Basel II” – are likely to promote the soundness and stability of the Swiss financial system.

The prospects for indirect market discipline are less promising. Even when market data were informative for the SFBC – which, however, does not appear to be the case – the Swiss supervisor should be very careful in using market signals as triggers for prompt corrective actions, as tying intervention to market prices may distort their informational content. Given today's level of knowledge, market indicators should simply represent a complementary source of information to be considered in the supervisory process. At the most, a negative market signal should prompt more intensive monitoring by the SFBC.

2 The prospects of market discipline for banks.

Evidence from the Swiss bond market

In recent decades, attention both among supervisors and in the academic community has focused increasingly on the role of the market in disciplining firms, and banks in particular. The new Basel Capital Accord (“Basel II”) explicitly considers market discipline as one of three pillars designed to improve the stability of the global financial system. There are two main reasons for this attention. First, the banking business is becoming more and more complex, especially at large (and systemic relevant) institutions. Because of this sophistication it is increasingly difficult for regulatory agencies to oversee banks in an effective way. Given that financial prices summarize the views and opinions about an institution held by thousands of investors and analysts around the globe, supervisors would wish to rely on the market in helping them to control and better evaluate the prospects of the institutions they monitor. Second, financial markets have evolved considerably in recent years. Today there is a wide range of products and securities for which the market is deep and liquid enough to allow regulators to extract accurate information from the data.

Typically, the literature differentiates between direct and indirect market discipline. Direct market discipline arises if there is a close relationship between the business policy of a firm and its funding costs that reduces the incentives for managers to behave improperly. For example, a bank may refrain from increasing its asset risk above efficient levels because otherwise the return required by depositors or other creditors would rise excessively, making debt too expensive. Conversely, researchers say that the market can indirectly discipline a bank when the information embedded in market data allows supervisors to better assess the prospects of that institution.¹

This paper has two main goals. First, I study the scope for direct market discipline in the Swiss banking sector by testing the validity of one of its main prerequisites, namely that the price of a financial security contains timely information about the risk profile of its issuer (market monitoring hypothesis). In a way, this represents a test of market efficiency for Swiss banks. Second, I assess whether those financial data may be useful to Swiss supervisors (indirect market discipline hypothesis). I do all this by analyzing the information content of bond spreads in the secondary market. The motivation for this choice mainly lies in the specific features of the Swiss financial services industry. If one excludes UBS and Credit Suisse Group (CSG), Swiss banks tend to issue only “traditional” securities, like stocks or senior bonds. Few institutions, moreover, borrow more than two times per year on the capital market. Because of this modest debt issuance activity, one has to rely on a panel of secondary bond market data in order to have enough observations for the

¹See Sironi (2003) for a careful characterization of the market discipline concept.

2 *The prospects of market discipline for banks. Evidence from the Swiss bond market*

empirical analysis, although primary market information may actually be more appropriate: As emphasized by Sironi (2001), primary markets are usually more liquid. Some supervisors consider (subordinated) debt spreads as the best financial indicator currently available also from a more theoretical point of view. In fact, bondholders are, with some exceptions, exclusively concerned with the downside potential of a bank.² In this sense they are averse to risk. Equity-holders, on the other side, are potential beneficiaries of big positive outcomes and are protected by limited liability in the case of a severe downturn. Because of this nonlinear payoff profile they may have an excessive appetite for risk, as shown by Dewatripont and Tirole (1994). As managers are expected to act on shareholders' behalf, this excessive appetite for risk may drive the business policy of the whole bank. Supervisors, on the other hand, try to limit excessive risk-taking in order to enhance efficiency and prevent banks from going bankrupt. Hence, their concerns and objectives align with those of subordinated debt-holders. In principle, depositors and other senior debtors would also be mainly interested in the downside potential of a bank. However, those investors are often less exposed to losses and therefore have fewer incentives to accurately monitor banks' activities. For those reasons, too, the majority of the proposals aiming at improving market discipline advocate the implementation of mandatory subordinated debt policies.³

Empirical studies providing evidence of market monitoring in industrialized countries abound. Most of them focus on the US (see Flannery, 1998 for an overview of the early literature and DeYoung, Flannery, Lang, and Sorescu, 2001; Swidler and Wilcox, 2002; Fan, Haubrich, Ritchken, and Thomson, 2003; Krainer and Lopez, 2004; Krishnan, Ritchken, and Thomson, 2006; Curry, Elmer, and Fissel, 2007 for more recent research). Many new contributions examine European banks (Persson, 2002; Sironi, 2002, 2003; Gropp, Vesala, and Vulpes, 2006) or Japanese banks (Brewer, Genay, Hunter, and Kaufman, 2003; Imai, 2007). Some researchers (e.g., Berger, Davies, and Flannery, 2000; Evanoff and Wall, 2001; Gunther, Levonian, and Moore, 2001; Cannata and Quagliariello, 2005) go a step further and argue that supervisory information and market information are complementary sources of bank supervisory intelligence. In other words, indirect market discipline appears to be feasible. Yet, there is also evidence (Hancock and Kwast, 2001; Evanoff and Wall, 2002) that some caution is needed when interpreting market data: Liquidity premia and other frictions may distort the information content of market signals.

This contribution differs from the existing literature in several aspects. First, the analysis is restricted to the Swiss banking sector. To my knowledge, empirical work on market discipline in Switzerland is limited to Birchler and Maechler (2002), who look at deposit-based funding and find that uninsured depositors tend to withdraw money from banks that experience a deterioration in their risk profile. In particular, nothing has been done on the role of the capital market, which, because of the sophistication of its participants, may be particularly relevant in this context. The Swiss case is quite interesting since a feature of the regulatory system is that it is mainly based on indirect supervision. With the exception of the two big banks UBS and CSG, the national supervisor

²In extreme situations, however – for instance when the bank is insolvent – this may not be true.

³See Board of Governors of the Federal Reserve System (1999) for an overview.

– the Swiss Federal Banking Commission (SFBC) – does not obtain internal information about a bank directly, but through an auditor. Other countries rely more on direct supervision. In the US, for instance, many more banks are subject to regular on-site inspections. One can then argue that the information thus obtained by US supervisors may be superior to the data available to the SFBC. If this were true, then in Switzerland the market could play a more important role as an information source and as a disciplining device than in other countries. Moreover, in view of the Enron and WorldCom cases, the market may be useful in disciplining the auditors as well, since the SFBC would be able to compare auditors’ reports with the current market opinion and eventually find out when the auditors lose the critical distance to their customers.

Second, in the literature mentioned, only large banks were usually taken into account. Given that the Swiss banking sector consists of UBS, CSG and about 330 smaller banks, I also consider the impact of market discipline on other bank groups. Of course, including small institutions in the sample brings about some additional problems, in particular concerning bond market liquidity, but it also raises interesting issues. As smaller banks represent a minor threat to systemic stability as opposed to big players, one would expect the monitoring of those banks to be less comprehensive. In such cases the market may be a particularly valuable and cheap information supplier for the supervisor. On the other hand, analysts and investors also tend to follow smaller banks less closely, as global asset and fund managers usually put their money in blue chips. Moreover, smaller institutions have a less transparent reporting policy compared to large banks. This potentially reduces the effectiveness of market discipline.

The explicit use of a model that tries to describe the relationship between risk and bond prices is another particular feature of this paper. This idea is obviously not new. Examples that illustrate the dangers in estimating “ad hoc” empirical models abound in the economic literature. Gorton and Santomero (1990) emphasize the importance of founding empirical tests regarding market discipline on a solid theoretical framework. The main feature of the model presented here is that it evaluates the interactions between all the parties concerned with the risk profile of a bank, namely the shareholders/managers and various types of debt-holders. In particular, I endogenize the risk decision process of the bank managers. Since other researchers considered only the relationship between the risk profile of a bank and bond spreads, leaving the risk choice unexplained, my paradigm gives some additional insights when interpreting the estimation results.

The empirical analysis confirms the international evidence that bond yields depend on the risk profile of the issuer. Investors are usually efficient and forward-looking in pricing risk. Cantonal banks that enjoy an explicit state guarantee are an exception, however. Investors appear to be less meticulous in pricing default risk for debt issued by these institutions. It is natural to conclude that bondholders believe that, thanks to the state guarantee, their investment is at no (or less) risk.

With respect to indirect market discipline, results are mixed. Although prices are risk-sensitive and quickly incorporate new information, it is not clear whether the market anticipates the supervisors in identifying problem at banks. While the data give little support for the indirect market

discipline hypothesis, one cannot exclude the possibility that this lack of evidence is due to the fact that Swiss regulatory bodies already use market information.

This paper is organized as follows. In Section 2.1 the theoretical model is presented and discussed. From this basis, I derive the empirical framework (Section 2.2). Section 2.3 discusses the results of the estimations and several robustness tests. Section 2.4 concludes.

2.1 The theoretical model

The main goal of this paper is to determine the feasibility of financial market discipline for Swiss banking organizations. Before trying to address this question empirically, I develop a theoretical framework in order to understand how the level of banks' endogenous risk⁴ is actually determined and how the price of debt is related to the risk policy chosen by banks. The analysis is built on a structural approach similar to the one formulated by Merton (1974). In his fundamental paper, Merton showed how equity and debt can be modeled as contingent claims on bank assets. Since its conception, Merton's structural approach has been widely used in analyzing market discipline and bond prices (see, e.g., Longstaff and Schwartz, 1995; Campbell and Taksler, 2003; Gropp, Vesala, and Vulpes, 2006). Several simplifying assumptions are needed in order to get a tractable model, though. Basic Merton models were often criticized because of the unrealistic features they are based on. This criticism is founded on their relatively poor empirical performance in pricing securities (Jones, Mason, and Rosenfeld, 1984).⁵ My paradigm probably suffers from the same drawbacks. Yet, I use the structural approach in order to understand the behavior of the investors involved, not to value bonds or stocks. Thus, I believe that the following basic assumptions will not significantly bias the analysis.

Imagine an economy with three agents: the shareholder/manager of a bank, a senior and a junior debt investor. Everybody is sequentially rational and risk-neutral. At the beginning of each period the bank has the opportunity to invest its assets A_{t-1} in a risky project. The gross return on the invested assets r_t^A is random, with $r_t^A \sim d[\mu(risk_t, \theta_t), \sigma(risk_t, \theta_t)]$. $risk_t$ is the level of endogenous asset risk chosen by the manager – i.e., there are both relatively safe and more risky projects to invest in – and θ_t is a set of other variables affecting mean μ and volatility σ of the return on assets. Higher moments are not taken into consideration, i.e., I implicitly make the simplifying assumption that preferences depend only on those two parameters.

In order to carry out this kind of business the bank needs sufficient funds. I assume that equity is

⁴I define endogenous risk as the component of default risk under the control of the bank's management. Total default risk is also affected by exogenous or common factors such as the interest rate or the cycle. Nowadays, however, the dichotomy has become less clear-cut, as financial instruments make it possible to hedge or speculate against almost every type of risk.

⁵As illustrated by Arora, Bohn, and Zhu (2005), however, modified structural models like the Vasicek-Kealhofer model (see Kealhofer, 2003) appear to be quite accurate in discriminating between defaulters and non-defaulters, also when compared to reduced form models. In addition, Morris and Shin (2004) have shown that after controlling for coordination problems between creditors, some empirical anomalies of the Merton model disappear.

2.1 The theoretical model

scarce and expensive so that the bank has to issue two types of debt: senior and subordinated one-period bonds. The manager proposes a contract to both types of debt investors that specifies a fixed one-period return for each security, r_t^S and r_t^J . The liquidation value of each contract at the end of that period t is given by S_t and J_t respectively. If either investor does not agree to finance the bank – that is, if one of them refuses to lend money to the bank at rate r_t^S and r_t^J respectively – everybody receives a risk-free return of r_f in that period, with $r_f > 1$. If indifferent, the investors will sign the contract. For technical reasons I further assume that the amount of senior and subordinated bonds needed at the beginning of each period t is a constant proportion (λ_S respectively λ_J) of the assets that should be invested. If all the contracts are signed, the manager of the firm chooses $risk_t$ to maximize the value of equity C_t at the end of period t . The bank defaults if, at the end of any period, the assets are less than total debt (face value of debt plus granted interests). There is no bail-out. The senior bondholder receives the minimum between $r_t^S \lambda_S A_{t-1}$ and $A_t = r_t^A A_{t-1}$, i.e., the value of the assets at the end of period t . In the event of bankruptcy, the subordinated debt-holder receives the value (if any) of the assets left after the senior investor has satisfied his claims. In other words, I assume a strict priority rule in the case of insolvency. If the bank does not default, the junior bondholder receives $r_t^J \lambda_J A_{t-1}$. The shareholder has a positive payoff only if the bank does not go bankrupt. In the absence of bankruptcy the shareholder can either liquidate the bank immediately or keep it in business. Finally, there are neither taxes nor bankruptcy costs and all the above is common knowledge.

Given those assumptions the following restrictions must hold at the end of every period t :

$$\begin{aligned} S_t &= \min [A_t, r_t^S \lambda_S A_{t-1}] \\ J_t &= \max [0, A_t - r_t^S \lambda_S A_{t-1}] + \min [0, r_t^S \lambda_S A_{t-1} + r_t^J \lambda_J A_{t-1} - A_t] \\ C_t &= \max [0, A_t - (r_t^S \lambda_S + r_t^J \lambda_J) A_{t-1}, V_t] \end{aligned}$$

where V_t is the present value of all present and potential expected future dividends for the shareholder.

The inclusion of two debt types may look like a futile complication of the model. Yet, subordinated debt has been often considered a better disciplining instrument in both literature and practice. Moreover, this liability structure is typical for financial intermediaries, as depositors are usually senior to all other creditors, bondholders in particular. From an empirical perspective, it is therefore important to make this distinction.

Some other assumptions may even appear unrealistic. In particular, banks are usually constrained on the amount of equity, not of debt. Moreover, the structure of the debt is in reality endogenously determined. However, managers normally have a preference for debt as a source of funds, since it is expected to be less expensive than equity. In addition, this assumption holds automatically in the context of a mandatory subordinated debt policy. That the liability side of banks' balance sheets usually does not change very much over time makes my assumption empirically founded.

2 The prospects of market discipline for banks. Evidence from the Swiss bond market

The strict priority rule and the “no bail-out” assumption are also questionable, given the discretionary power of regulators in deciding whether to keep a troubled bank open or not or whether to restructure it. I will return to this point in Section 2.2.

Gorton and Santomero (1990) employ a very similar approach. Their model is based on the Black and Cox (1976) multiple debt claimants version of the standard contingent claims valuation framework developed by Black and Scholes (1973). A feature of those models is that they assume log-normality of the asset return. However, Gorton and Santomero did not specify the risk decision of the bank; that is, the level of risk chosen by the manager is an exogenous variable. By endogenizing this aspect I have to rely on simpler distributions, since the normality assumption would make the model intractable.

2.1.1 One-period game

Consider first a one-period game. Suppose that the asset return is uniformly distributed with mean μ_σ , which may depend on the risk chosen by the manager, and parameter σ , a measure of the variability of the return on assets (the standard deviation of the distribution equals $\frac{\sigma}{\sqrt{3}}$). The higher σ , the riskier the business run by the bank. Note that in this setting bank default risk is fully captured by σ . For simplicity, I normalize the value of the assets at the beginning of the period to 1. Those assumptions imply that

$$\begin{aligned} A &= r^A \sim U[\mu_\sigma - \sigma, \mu_\sigma + \sigma] \\ f(A) &= \frac{1}{2\sigma} \text{ and } F(A) = \frac{A - \mu_\sigma + \sigma}{2\sigma}. \end{aligned}$$

$f()$ and $F()$ are respectively the pdf and cdf of the future value of the assets. In addition $\sigma \in [\sigma_{\min}; \sigma_{\max}]$, with $\sigma_{\max} \leq \mu_\sigma$ and $\sigma_{\min} > \mu_\sigma - r_f \lambda_S$. The first inequality ensures that the value of the assets is always nonnegative. This constraint does not seem unrealistic as bank regulation and oversight should limit excessive risk-taking (see Stever, 2007). Besides regulatory constraints, factors like reputational or legal risk may also limit the risk appetite of the manager. The second restriction has two major implications. First, the asset value at the end of the period can fall below $r_f \lambda_S$. In other words, senior bonds are risky securities as there is always a probability that ex post the senior debt-holders may regret having invested in the bank. Secondly, the banking business cannot be risk-free and at the same time attain a “free lunch” return μ_σ , which I assume is always greater than the risk-free rate. This last assumption makes economic sense – under normal circumstances it would be unreasonable to invest in a project whose expected return is lower than the risk-free rate – and allows a neat derivation of the equilibrium constraints.⁶ In particular, it is straightforward to show that if the expected return on assets is high enough when compared to the risk-free rate, in equilibrium μ_σ always exceeds the funding costs per unit of invested assets $r^S \lambda_S + r^J \lambda_J = \tilde{\lambda}$. $\mu_\sigma > \tilde{\lambda}$, in turn, is a sufficient condition for obtaining a reasonable reduced-form

⁶See Appendix 2.5.1 and Appendix 2.5.2.

2.1 The theoretical model

equilibrium. Note that $V = A - \tilde{\lambda}$ in this one-period case, so that the restriction for equity reduces to $C = \max [0, A - \tilde{\lambda}]$.

The manager maximizes shareholder value by choosing an optimal level of σ , which I call σ^* . In technical terms the manager solves

$$\max_{\sigma} E(C) \quad (2.1)$$

$$\text{s.t. } E(C) \geq (1 - \lambda_S - \lambda_J) r_f. \quad (2.2)$$

Constraint (2.2) states that the equity-holder will not liquidate the bank at the beginning of the game only if the bank's expected profits cover the opportunity costs of equity. The expected value of equity at the end of the game is given by

$$\begin{aligned} E(C) &= E(A - \tilde{\lambda} \mid A > \tilde{\lambda}) \text{prob}(A > \tilde{\lambda}) \\ &= \left(\frac{\mu_{\sigma} + \sigma + \tilde{\lambda}}{2} - \tilde{\lambda} \right) \left(\frac{\mu_{\sigma} + \sigma - \tilde{\lambda}}{2\sigma} \right) \\ &= \frac{(\mu_{\sigma} + \sigma - \tilde{\lambda})^2}{4\sigma}. \end{aligned} \quad (2.3)$$

Suppose for a moment that the expected return on assets does not depend on the level of risk, i.e., $\mu_{\sigma} = \mu$. Then differentiating (2.3) with respect to σ one gets

$$\frac{\partial E(C)}{\partial \sigma} = \frac{\sigma^2 - (\mu - \tilde{\lambda})^2}{4\sigma^2}$$

which is always positive for $\sigma \in [\sigma_{\min}; \sigma_{\max}]$, since, by assumption, $\mu > \tilde{\lambda}$ and $\sigma_{\min} > \mu - r_f \lambda_S > \mu - \tilde{\lambda}$. In other words, the manager will always choose the highest possible level of σ independently of the borrowing costs. This is what one would have expected, since in this simple one-period game equity can be viewed as a European call option on the bank's assets. And option values increase with the volatility of the underlying security, which in this case is the value of the assets.

To arrive at an internal solution, the risk appetite of the manager must be made concave, for instance by choosing an alternative functional form for μ_{σ} . Suppose for instance that

$$\mu_{\sigma} = \mu_M - \kappa (\sigma - \sigma_M)^2 \quad (2.4)$$

where μ_M is the maximum expected return that the bank can achieve by choosing the appropriate level of risk σ_M . κ is a parameter that determines the degree of dependence between the return on assets and risk. In Appendix 2.5.3, I show that for a given value of κ an internal optimum may indeed exist. But how can one justify such a functional form for μ_{σ} ? That the expected return on assets increases with σ for low levels of risk is consistent with basic portfolio theory. Reputational

2 The prospects of market discipline for banks. Evidence from the Swiss bond market

or legal issues may provide a rationale for the negative relationship between risk and return for high values of σ .

Independently of the functional form of μ_σ , $E(C | \sigma)$ depends on $\tilde{\lambda}$. Thus, in order to determine the subgame perfect Nash equilibrium of the game one needs to know what debt investors will do. Note that, because of the common knowledge assumption, both lender types know what level of risk σ^* the manager will choose if the contracts are signed.

Consider first the senior debt-holder. Under risk-neutrality he will accept the deal when

$$E(S | \sigma^*) \geq r_f \lambda_S.$$

I show in Appendix 2.5.1 that this condition is satisfied if

$$r^S \geq \frac{\mu_{\sigma^*} + \sigma^* - [4\sigma^* (\mu_{\sigma^*} - r_f \lambda_S)]^{\frac{1}{2}}}{\lambda_S}. \quad (2.5)$$

It is straightforward to demonstrate that $\frac{\partial r^S}{\partial \sigma^*} > 0$ for $\sigma \in [\sigma_{\min}; \sigma_{\max}]$ and $\mu_\sigma = \mu$.⁷ Hence, the required bond return increases with asset risk, as one would have expected.

For the subordinated debt-holder the problem is very similar. She will sign the contract if

$$E(J | \sigma^*) \geq r_f \lambda_J.$$

Given the equilibrium value of r^S , this holds when

$$r^J \geq \frac{[4\sigma^* (\mu_{\sigma^*} - r_f \lambda_S)]^{\frac{1}{2}} - [4\sigma^* (\mu_{\sigma^*} - r_f \lambda_S - r_f \lambda_J)]^{\frac{1}{2}}}{\lambda_J}. \quad (2.6)$$

In equilibrium (2.5) and (2.6) hold with equality, since I assumed that the manager has all the bargaining power is setting the terms of the contract. Together with (2.1) they characterize the subgame perfect Nash equilibrium. As an example imagine that $\mu = 1.2$, $r_f = 1.01$, $\lambda_S = 0.5$ and $\lambda_J = 0$. Then the following conditions must hold for a subgame perfect equilibrium in which the bank can carry out its business: (i) $\sigma \leq 1.2$, (ii) $\sigma > 0.695$ and (iii) $r^S \geq \frac{1.2 + \sigma - 2\sqrt{0.695\sigma}}{0.5}$. Because the manager maximizes risk and expected value of equity, (i) and (iii) are binding, that is, $\sigma = 1.2$ and $r^S = 1.15$. Note that $E(C) = 0.7$, so that the shareholder strictly prefers to run the bank than to invest his money in risk-free securities.

2.1.2 Infinitely repeated game

Consider now an infinitely repeated game. I assume that the distribution of r_t^A is the same as before in every period and that at the end of each period one can have the following three situations:

- a) $A_t < (r_t^S \lambda_S + r_t^J \lambda_J) A_{t-1}$. In this case the bank defaults and the shareholder gets nothing.

⁷See Appendix 2.5.1. When $\frac{\partial \mu_\sigma}{\partial \sigma}$ is high, however, this may not be true.

2.1 The theoretical model

- b) $A_t - (r_t^S \lambda_S + r_t^J \lambda_J) A_{t-1} > A_{t-1} - (\lambda_S + \lambda_J) A_{t-1}$. The equity value grew between t and $t-1$. In this case the shareholder receives a cash dividend D_t that corresponds to the increase in the value of book equity, i.e., $D_t = A_t - [1 + (r_t^S - 1) \lambda_S + (r_t^J - 1) \lambda_J] A_{t-1} > 0$.
- c) $0 < A_t - (r_t^S \lambda_S + r_t^J \lambda_J) A_{t-1} < A_{t-1} - (\lambda_S + \lambda_J) A_{t-1}$. The bank business generated losses, but the bank remains solvent. In this case the shareholder has two options. He can either shut down the bank and receive the residual $A_t - (r_t^S \lambda_S + r_t^J \lambda_J) A_{t-1}$ or bring additional capital $D_t = A_t - [1 + (r_t^S - 1) \lambda_S + (r_t^J - 1) \lambda_J] A_{t-1} < 0$, and leave the bank in business.

Those assumptions have several implications. First, if the bank is still in business, the value of the assets at the beginning of each period is constant and equal to $A_0 = \bar{A}$. The amount of senior and subordinated debt needed by the bank does not vary either, and equals $\lambda_S \bar{A}$ and $\lambda_J \bar{A}$ respectively. Second, since the game faced by the players is the same at every period (stationary game), the optimal values of σ , r^S and r^J , which have to be determined, also remain constant over time. Finally, those assumptions imply that $A_t \mid \bar{A} = r_t^A \bar{A} \sim U[(\mu_\sigma - \sigma) \bar{A}, (\mu_\sigma + \sigma) \bar{A}]$. As above, I set $\bar{A} = 1$ without loss of generality.

If the game reaches the situation described in c), the shareholder will always choose to bring the additional capital and keep the bank in business. To see this, note first that in this case I know that at time $t = 0$ the shareholder chose to run the bank, which means that the expected value of operating the bank exceeded the expected profit of having $1 - \lambda_S - \lambda_J$ in cash. Second, in state c) the shareholder is almost in the same situation. The a priori distribution of the asset value is always identical and therefore expected benefits from keeping the bank in business are the same as in $t = 0$ less the required additional capital. The shareholder does not shut down the bank when those profits are greater than the benefits of having $A_t - r^S \lambda_S - r^J \lambda_J$ in cash. Those two conditions can be written as

$$\begin{aligned} V_0 &\geq 1 - \lambda_S - \lambda_J \\ V_t + A_t - 1 - (r^S - 1) \lambda_S - (r^J - 1) \lambda_J &\geq A_t - r^S \lambda_S - r^J \lambda_J \end{aligned} \quad (2.7)$$

where V_t are the expected benefits of keeping the bank in business at time t . (2.7) reduces to

$$V_t \geq 1 - \lambda_S - \lambda_J$$

But $V_t = V_0$ in such a stationary game so that (2.7) must also be satisfied, if the shareholder already chose to run the bank at $t = 0$. It follows that if the bank did not default at the end of period $t - 1$, the risk-neutral manager will maximize

$$E_{t-1}(C_t) = E_{t-1} \left(\sum_{i=0}^{\infty} \frac{D_{t+i}}{r_f^i} \right)$$

over σ . In Appendix 2.5.2 it is proved that

$$E_{t-1}(C_t) = \frac{r_f \delta(\sigma) \pi(\sigma)}{r_f - \pi(\sigma)} \quad (2.8)$$

with

$$\begin{aligned} \delta(\sigma) &= \frac{\mu_\sigma + \sigma - \tilde{\lambda}}{2} - (1 - \lambda_S - \lambda_J) \\ \pi(\sigma) &= \frac{\mu_\sigma + \sigma - \tilde{\lambda}}{2\sigma}. \end{aligned}$$

$\delta(\sigma)$ represents the expected value of the dividend in each period given no default, while $\pi(\sigma)$ is the probability of receiving this dividend in that period.

To find the level of risk chosen by the manager one has to maximize (2.8) over σ at $t = 0$. Again, suppose first that μ_σ is independent of σ . In Appendix 2.5.2 it is also shown that for $\sigma \in [\sigma_{\min}; \sigma_{\max}]$ and realistic parameter values (2.8) is a convex function of σ .⁸ That is, the manager will choose either a very high or a very low level of risk. How can one explain this outcome as opposed to case 1? While equity can still be viewed as a call option on the bank's assets, this option is not the same as before. In fact, when the underlying is greater than the strike price at maturity in period t (that is if $A_t > \tilde{\lambda}$), the shareholder receives not only a cash dividend but also another identical call option. By increasing volatility, the manager still raises the value of the expected value of the dividend, but at the same time he reduces the probability of receiving the next option and all the others that may follow. In other words: in contrast to the previous case, the bank has a franchise value that depends on σ . So there is a trade-off: The manager either chooses to maximize the value of the limited liability component of equity or try to minimize the risk of losing the franchise value of the bank. Accordingly, one has $\frac{\partial \delta(\sigma)}{\partial \sigma} = \frac{1}{2} > 0$ and $\frac{\partial \pi(\sigma)}{\partial \sigma} = -\frac{\mu - \tilde{\lambda}}{2\sigma^2} < 0$.

Again some concavity is needed in order to have an internal solution. For instance, I show in Appendix 2.5.3 that with the functional form for expected return on assets assumed in (2.4) one can have a local maximum. In this case there would be three "forces" that determine the risk choice of the manager. In addition to the limited liability and franchise value effects described above, the manager has an incentive to choose the level of risk so that the asset distribution is optimal in terms of assets' return.

For the debt-holders the same restrictions as in Section 2.1.1 apply.

2.1.3 Discussion

Equations (2.5), (2.6) and (2.8) provide several interesting insights that should be taken into consideration when implementing the empirical model and in interpreting its results. First, they illustrate how both equilibrium bond yields depend positively on the asset risk in a nonlinear way. Second,

⁸This is consistent with Park and Peristiani (2007), who empirically analyze how the put option value and the franchise value of a bank interact with risk and observe a convex nonlinear relationship between the market-to-book ratio and the risk of failure.

note that a change in one variable alters the whole equilibrium vector $\{\sigma; r^S; r^J\}$. Take for instance the expected return on assets μ . Appendix 2.5.1 shows that bond spreads are negatively correlated with μ . This looks intuitive, since if the expected return on the assets increases, substantial losses for the debt-holder would be less likely. Yet, in doing this kind of comparative static analysis, one keeps all other variables constant. But a change in μ affects the risk choice of the manager as well. This can be seen by looking at (2.8). Intuitively as μ rises, the franchise value of the bank increases and for the manager it is now optimal to choose a lower level of risk. Rational debt investors anticipate this and adjust the required risk premium accordingly. This indirect effect of μ on the spread is absent in models that do not endogenize the decision process of the manager. Strictly speaking, however, the convexity of (2.8) induces the manager either to minimize or to maximize risk and hence comparative static becomes less useful. One can overcome this problem by assuming (2.4), which may lead to an internal optimum.

Even in this case the model in the form presented above is too simple to be used for empirical work. Since there is no asymmetric information between manager and debt-holders and the value of the assets in one period is the only stochastic variable, the manager will always choose the same volatility level, and the debt-holders will always accept or refuse the contract. This is of course very unrealistic. Managers are usually better informed than the investors, especially concerning the risk-return trade-off. In other words, not all the variables are constants known in advance and the manager has a better view on how the bank business develops. For example, assume that in my simplest model – where the return on assets is independent of σ – μ_t is a martingale and that the manager knows the current value of μ_t at the beginning of each period t whereas the debt-holders can only infer the value of μ_t by observing the value of the assets at the end of that period, that is after having signed the contract. While the maximization problem remains the same, the debt-holders must in this case be aware of the fact that they do not know in advance what level of risk will be chosen, since σ_t is a function of μ_t . In equilibrium (if any) they can only guess what the asset return will be and eventually adjust their estimate in the next period. In this more plausible setting the optimal value of σ_t may change over time.

The simplistic assumptions underlying the distribution of the return on invested assets is another unrealistic feature of the model presented above. First, there are a lot of exogenous factors like the business cycle, the interest rate, and so on that can affect this distribution. Second, the uniformity hypothesis is clearly too simplistic. Skewed distributions, such as the lognormal distribution, would surely be better at describing the reality.

The return on debt also depends on many variables and not only on the default risk premium. Some of those factors, like the risk-free term structure, priority rules and recovery rates, are already incorporated in my model. Others, like bankruptcy costs and taxes, could easily be included. However, I believe that the inclusion of bankruptcy costs and taxes would not add more value in this context. If I assume that bankruptcy costs and taxes are constant over time, this would probably affect the general level rather than the dynamics of the equilibrium, which is what I am actually interested in when analyzing market discipline empirically. Moreover, in Switzerland

corporate and government bonds are subject to the same taxation rates and therefore the tax effects on the spread should mainly cancel out.⁹ Finally, my model ignores some factors that could be even more important than default event risk in determining the evolution of bond yields. These factors include liquidity premia, common risk or the general degree of risk aversion. In implementing the econometric model, I have to address all of those issues.

2.2 Empirical analysis

The theoretical model of Section 2.1 shows how banks choose asset risk and how this should be reflected in bond prices in the presence of market discipline.¹⁰ The next step is to test whether this relationship also exists in the real world. Accordingly, the objective of the next two sections is to find out whether secondary market bank-bond spreads contain information about the idiosyncratic asset risk profile of the issuer and when adjustments to new information occur (market monitoring hypothesis). In addition, I try to establish whether the information (if any) embedded in debenture prices may be useful to supervisors (indirect market discipline hypothesis). The model outlined in the previous section gives useful insights in this context. In particular, it highlights other important determinants of bond spreads besides asset risk, like for instance the structure of the liability side of the balance sheet. Moreover, the model illustrates how some variables can have an indirect effect on the spread, by affecting the risk choice of the manager. In interpreting the empirical results, it is important to be aware of those interactions.

Sironi (2001) argues that even for large banking organizations the secondary market in Europe is not liquid and developed enough and therefore does not allow us to extract accurate information from the data. Only issuance spreads appear to accurately reflect the riskiness of banks. This statement may be particularly true in Switzerland: A large share of debt is purchased by institutional investors, who usually hold those securities until maturity. It is very possible that for some bonds no trading takes place for several days. In such cases the indicative prices set by the market makers may also not be very informative. In light of those facts, it seems futile to seek evidence of indirect market discipline in the Swiss secondary bond market. Some points are worth noting, though. Even if the secondary bond market is not as liquid as, for example, the secondary stock market, some trading *does* take place. Moreover, institutional investors are professional, constantly well-informed investors. If there is room for arbitrage gains – i.e., if they believe one particular bond is clearly mispriced after observing the last trade and this compensates the expected liquidity

⁹There is empirical evidence that taxes are an important determinant of bond spreads. See, e.g., Elton, Gruber, Agrawal, and Mann (2001) and Gabbi and Sironi (2005). However, it is important to point out that in some countries like the US corporate bonds and treasuries may be subject to different tax rates.

¹⁰Note that direct market discipline in the narrow sense is absent in my model, as the bank chooses asset risk after the contracts are signed and there is no asymmetric information between manager and debt-holders. By introducing asymmetric information in a multiperiod game – e.g., with respect to μ_t – it would be possible to analyze the impact of market discipline on the risk choice of the manager. This is, however, beyond the scope of this paper.

premium – they will probably seize the opportunity and sell or buy the bond. Therefore, some information must flow into secondary market prices too.

Secondary market data also have some technical advantages, both in the context of my study but also concerning their use as early warning indicators. First, and most importantly, the data set would be extremely small, if one were to focus the analysis on the Swiss primary market. For example, only 7 bank holding companies (13 banks) had subordinated debt outstanding in 2001. Furthermore, few of them borrow on the capital market more than twice a year. By relying on a panel of secondary market data, I am able to obtain a sufficient number of observations. Second, the decision of whether to issue senior or subordinated debt has per se some information content which is embedded in primary market data (Birchler and Hancock, 2004). For example, a bank may decide to issue subordinated debt instead of senior debt in order to improve its Tier 2 or Tier 3 capital. By focusing on price changes in secondary market data I avoid dealing with these issues in my analysis. Finally, secondary market spreads can easily be used to assess how the risk profile of a bank evolves over time.

2.2.1 Variable description

One of the issues described in Section 2.1.3 is that bond spreads depend on several factors, some of which may be unrelated to the risk profile of the issuing bank. In other words, although I am mainly interested in asset risk, it is important to control for all determinants of bond spreads. In the following I describe all the variables which I include in the estimations.¹¹

Bond spread

All the information concerning the single bond series, including prices and internal rates of returns, are obtained from Thomson Datastream and Bloomberg Professional. Price updates reflect the last official trade at the Swiss Exchange in Zurich. The spread ($sp_{i,t}$) is simply the difference between the redemption yield of a bank debenture and that of the Swiss government bond with the closest maturity. I assume that government debentures are risk-free and liquid enough such that neither default risk nor liquidity premia must be taken into account. I keep callable or other exotic securities out of the sample. Consequently, the internal rates of return need not to be adjusted in order to control for embedded options. Observations are discarded whenever time to maturity is less than one year, since as a bond approaches maturity its redemption yield becomes very volatile and may begin to behave abnormally.

¹¹A table with a short description of the variables used and their availability can be found in Appendix 2.5.4. Unless otherwise stated, subscript b stands for a bank-specific variable, while subscript i stands for a bond-specific variable.

“Observable” bank-specific asset risk

As stated above, my study focuses on the relationship between bond spreads and the risk profile of a bank. More definitions of risk are possible, depending on the purpose of the analysis. One can either focus on total risk or on endogenous risk. Direct market discipline can only affect endogenous risk by definition, while changes in security prices should also reflect changes in common risk. However, it is extremely difficult to find adequate measures of either type of risk. Volatilities are often used as a proxy for the total risk of a particular security. In my case the underlying security are the assets of a bank. Unfortunately the asset value is not directly observable. One could employ Merton’s model again in order to calculate the implied volatilities of the firm value (Crosbie and Bohn, 2002). The problem is that this methodology relies heavily on financial (usually stock) prices. If the equity market contains the same information as the bond market and this information does not only reflect asset risk (which is very likely – see below), using asset volatility as variable in my estimation may be problematic and lead to spurious results. Hence, I have to rely on other, more partial measures.

In many studies on market discipline, bank risk is proxied by ratios constructed from balance sheet data. I use several ratios obtained from the database of the Swiss National Bank. Following Morgan and Stiroh (2001) I differentiate between asset and performance variables. Clearly, in the banking business there are risky and less risky activities and assets. For instance, a residential real estate loan is considered a safer investment than a direct investment in a hedge fund. To measure this asset mix effect I employ the ratio of regulatory risk-weighted assets to total assets ($busrisk_{b,t}$). For banks active in relatively safe businesses this ratio – and, *ceteris paribus*, the spread – should be particularly low.

Because of the limitations in the current capital framework (“Basel I”), however, the risk weights do not say much about whether a bank has taken more or less risks *within* asset classes. Performance variables like the absolute level of provisions and credit losses to total credits ($provision_{b,t}$) or the amount of new provisions credit losses to total credits built up in the reference year ($newprov_{b,t}$) should be able to capture those nuances, since they reflect all the current and expected losses on the (credit) portfolio of that institution. Moreover, provisions and loan losses reduce the capital resources of the bank. Similar information is enclosed in the ratio of non-performing loans to total credits ($npl_{b,t}$) and in the return on assets ($roa_{b,t}$). However, the expected correlation between profitability and spread is *a priori* not clear. On the one hand, net losses reduce equity reserves and indicate that the bank has probably taken bad risks. On the other hand, above-average profits may also reflect an extremely aggressive strategy of that bank which turned out to be successful. Therefore, I expect the relationship between spread and profitability to be somewhat convex.

Finally, it is sometimes argued that liquidity problems may cause highly solvent companies to run into difficulties. In the banking business, the role of liquidity may be even more essential, given the short-term nature of deposits (Diamond and Dybvig, 1983). Hence, the amount of liquidity could be an important factor in determining the risk profile of a borrower. I measure bank liquidity

by the following quick ratio

$$liquidity_{b,t} = \frac{liquid\ assets_{b,t} - 33\% \text{ of short term liabilities}_{b,t}}{33\% \text{ of short term liabilities}_{b,t}},$$

which measures the liquidity surplus in percent of regulatory liquidity requirements.

Capital structure

The risk perception of investors should also be influenced by the capital structure of a bank (*deposit*_{b,t}, *seniord*_{b,t}, *juniord*_{b,t}). In general, the amount of capital (the leverage) should be negatively (positively) correlated with the spread. Bondholders see capital as a security buffer in the event of losses on the bank's assets. If those resources are high, the probability that the bank will default in the near future is low. Indeed, with a little algebra one can prove that for a given level of asset risk $\sigma \in [\sigma_{\min}; \sigma_{\max}]$, $\frac{\partial r^S}{\partial \lambda_S}$ and $\frac{\partial r^J}{\partial \lambda_J}$ are greater than 0.¹² In addition, the option owned by the equity-holders becomes relatively more valuable when the bank is highly leveraged, which may induce the managers to assume more risks. That a higher leverage and higher costs of debt induce borrowers to undertake riskier projects is a standard result (see for example Stiglitz and Weiss, 1981). For the static model it is straightforward to show that $\frac{\partial^2 E(C_1)}{\partial \sigma \partial \lambda} > 0$, that is, increasing risk becomes more attractive when $\tilde{\lambda}$ is high. In a dynamic context, things are more complicated and I cannot prove analytically that this is true. However, note that $\frac{\partial^2 \delta(\sigma)}{\partial \sigma \partial \lambda} = 0$ and $\frac{\partial^2 \pi(\sigma)}{\partial \sigma \partial \lambda} > 0$. In other words, while the marginal increase in the dividend when the bank assumes more risks is not affected by the leverage, the probability of losing this dividend decreases less if the bank has more debt. Moreover, in the Black-Scholes model the vega of an in-the-money call option, which can be interpreted as shareholders' marginal utility of augmenting risk when the bank is solvent, increases if the strike price, i.e., total liabilities, goes up.

Besides the leverage, other specific features of the capital structure affects the bond spreads, as shown in equations (2.5) and (2.6). If priority rules are strictly satisfied, senior debt investors regard capital and subordinated debt as equivalents in the case of bankruptcy. Therefore, the spread of senior bonds should depend negatively on the amount of subordinated debt issued for a given amount of equity capital. For the same reason, subordinated investors would require a higher spread if, all else being equal, a bank has a lot of senior debentures outstanding. Note that, in contrast to the previous section, all those "sources of risk" are related to the liabilities side of the balance sheet.

There is a problem that can bias the empirical analysis, though. Suppose that the probability of default of a bank (and therefore the spread) suddenly increase for some reason. The supervisor may then require the bank to increase its capital or to reduce its leverage. If one cannot perfectly control for the increase in bankruptcy risk, a regression may suggest that capital and spread are positively correlated.

¹²See Appendix 2.5.1.

“Unobservable” bank-specific risk

The SFBC assigns an internal rating to individual organizations which is comparable to the CAMEL or BOPEC rating assigned by the US supervisor to resident banks and bank holding companies respectively. In addition, an institution is put on a “black list” ($ebk_{b,t}$ equals 1), if the Swiss regulator believes that the bank has significant unresolved problems. It is important to point out that those unresolved problems may not be directly related to asset risk. Instead, this indicator reflects a broader opinion of the SFBC about the soundness of a bank. I assume that $ebk_{b,t}$ incorporates all information relevant to the Swiss supervisor. Accordingly, $ebk_{b,t}$ may be affected by factors like the management of the bank, its organizational structure or other possible sources of concern, for which I do not have specific information at my disposal. In contrast to the variables described above, this “black list” is strictly confidential – i.e., it cannot be observed by investors. Thus, a significant positive coefficient would signal that market participants are highly skilled and can identify sources of problems which are not accounted for by the standard variables described before.

General business and economic environment

Like every branch of the economy, the banking business is affected by exogenous factors. For instance, Lehmann and Manz (2006) show that the Swiss banking sector is exposed to fluctuations in the real economy. This is especially true for mortgage or credit banks, as during a recession the number of defaults generally increases. Typical indicators for the business cycle are the real GDP growth rate ($gdpgr_t$) or the unemployment rate ($unempl_t$). Although the business cycle represents a common source of risk for the whole sector and is beyond the control of the managers, the exposure in the credit or mortgage market may vary across banks. Even institutions with similar business rationales may have different exposures, as new instruments such as credit derivatives have increased the flexibility in managing credit risk. In other words, the risk premium related to the macroeconomic cycle should be neither a pure common nor a pure endogenous (bank-specific) driving factor of bond spreads. Yet, for simplicity, in the estimations I assume that the influence of the cycle on bond spreads is the same for all banks.

Another proxy for the general economic conditions is the risk-free rate ($riskfree_t$). The net effect of the interest rate on bond spread is a priori ambiguous. During recessions one often observes falling rates, at least at the left end of the yield curve. In the long run, however, a drop in the risk-free rate should stimulate the economy, and reduce the risk of default. Most importantly, interest rates also affect the risk choice of the bank managers. In other words, an increase in r_f reduces the franchise value of the bank, since future profits have to be discounted at a higher rate, while debt investors require higher compensation in absolute terms for the money they lend to the bank, which reduces future profits even further. Hence, the option component of equity becomes relatively more important and the managers may be willing to increase risk. Finally, in every standard structural model an increase in the risk-free rate implies an increase in the expected

growth rate of the enterprise value, which lowers the probability of default. This effect is apparently absent in my model, as I did not model the dynamics of the asset value as an explicit function of r_f . Consider instead the role of μ . According to equations (2.5) and (2.6) this expected return on assets and bond spreads should be negatively correlated for a given σ , as shown in Appendix 2.5.1. The intuition behind this result is straightforward: The probability of the bank defaulting diminishes as μ increases and therefore debt investors are willing to accept a lower return for a given level of endogenous risk. For empirical analysis it is presumably not sufficient to proxy μ by the risk-free rate, as in a risk-neutral world. Since the business of a typical bank in my sample encompasses a wide range of on- and off-balance-sheet activities, I employ several factors to model μ . First, the interest rate differential for the whole Swiss banking sector (interest revenues minus interest expenses to total credits, ird_t) and macroeconomic indicators like those mentioned before may be good proxies for the profitability of the loan and mortgage business. Secondly, I assume that fee, commission and trading income depends primarily on the mood of financial markets. Hence, as proxy for the profitability of those activities I use the market excess return ($market_t$). Note that μ directly affects the risk decision of the bank's managers as well. If the expected return on assets increases, the franchise value of the bank rises. For the management, it is then optimal to choose a lower level of risk, which should reduce the spread even further.

Systematic risk

Basic financial theory implies that corporate debenture investors would require an additional premium for undiversifiable risk if bond returns move systematically with other assets in the market. Empirical studies (see, e.g., Collin-Dufresne, Goldstein, and Martin, 2001; Elton, Gruber, Agrawal, and Mann, 2001; Campbell and Taksler, 2003) show that a common systematic component is indeed one of the most important determinants of bond spreads. If common risk premia are exclusively related to default risk, the variables described in the previous section should already capture those effects, at least partially. However, there are several reasons that could explain such a premium even in the absence of a common default risk. For example, Elton, Gruber, Agrawal, and Mann argue that expected default losses on corporate bond may be directly correlated with stock market prices. Moreover, as investors need to be rewarded for the risk they assume, changes in the required compensation for risk can introduce correlations across asset categories. More generally, Collin-Dufresne, Goldstein, and Martin suggest that supply and demand shocks in financial markets that are independent of default risk are likely to be important drivers of credit spreads.

As I am interested in market discipline at the single-bank level, I should address those systematic effects in my empirical analysis.¹³ Following Fama and French (1993) I employ a multiple factor model. A term premium ($term_t$) and a default premium factor ($default_t$) should capture common effects in the debt market. The three stock market factors of Fama and French are the market

¹³Driessen (2005) found that, after controlling for the common factors, the risk associated with movements in the firm-specific factors is hardly priced into corporate bonds. However, he finds some evidence of a default jump risk premium.

2 The prospects of market discipline for banks. Evidence from the Swiss bond market

excess return factor, the small minus big cap factor (smb_t) and the high minus low book-to-market factor (hml_t). As Campbell and Taksler find that equity volatility can explain a significant part of corporate bond yield indices, I expand the Fama and French model by adding an implied volatility index ($volimpl_t$).

Fama and French showed that, except for low-grade corporations, the two bond market factors can explain a large part of the common variation in bond returns. In my framework those two factors should also account for an important share of the systematic component in bond premia. However, I employ the richer specification as the stock market variables may be correlated with default risk and other spread determinants. As I already pointed out, the expected fee and commission income of a bank, for instance, depends directly on the situation in the financial markets, which is proxied by the market excess return. As a corollary, one has to be careful in interpreting the coefficients on the factors.

Market liquidity and other bond-specific characteristics

I define the market liquidity premium as the short-term deviation between transaction prices and long term fundamental value of a security, which depends mainly on endogenous and common default risk and possibly on institutional factors like taxes. These differences arise even in a perfectly rational world because of frictions reflecting order handling or inventory costs (Stoll, 1978; Ho and Stoll, 1983), asymmetric information (Kyle, 1985; Glosten, 1994) or strategic supply (Kyle, 1989; Glosten, 1989; Biais, Martimort, and Rochet, 2000). In addition to distorting the absolute value of the bond spread, low liquidity also decreases its informational content. If nobody trades a particular debenture because of the liquidity premium, the price of this security just reflects the opinion of the market maker or dealer and not that of the whole market.

Hancock and Kwast (2001) as well as Covitz, Hancock, and Kwast (2002) provide evidence that the market liquidity premium is indeed an important component of bank debt spreads. Liquidity is in general relatively low in the Swiss bond market. As proxies for market liquidity I use the amount issued ($amount_i$) and relative time to maturity ($mtmr_{i,t}$)¹⁴. First, it seems reasonable to assume that for larger issues the market is broader and more liquid. Second, empirical research (Board of Governors of the Federal Reserve System, 1999) provides evidence that the older the bond, the less frequently it will be traded. Elderly issues become absorbed into investors' portfolios and held until maturity. This effect should be particularly important in Switzerland, where institutional investors like pension or mutual funds play a major role in the bond market. Indeed, for most of the bonds in my sample the turnover is high just after the issue date, but declines rapidly thereafter.

Note that I do not employ the absolute value of time to maturity ($mtm_{i,t}$) in order to assess the importance of liquidity risk. This is because the time to maturity has an effect on the spread which is not only related to liquidity, but also to default risk. To see this, consider two zero bank bonds that pay 1 in one and two years respectively. Suppose there is a probability $1 - \pi$ that

¹⁴I define relative time to maturity as the ratio of the residual time span until maturity to total lifetime. At the issuing date this ratio equals 1, at maturity 0.

the bank defaults every year. In the case of default, the bond becomes worthless. Moreover, the risk-free yield curve is flat, with $r_f = 1$ for simplicity and both the liquidity and the systematic risk premium equals zero. Investors are risk-averse with utility function of the form $U(x) = \ln(1+x)$, so that $U(1) = \ln 2$ and $U(0) = 0$. The competitive equilibrium price B_1 of the one-year bond is then equal to $2^\pi - 1$ while $B_2 = 2^{\pi^2} - 1$. The implied spread sp_i can be obtained by solving $B_i = \frac{1}{(r_f + sp_i)^i}$ for $i=1,2$. What one gets is $sp_1 = (2^\pi - 1)^{-1} - 1$ and $sp_2 = (2^{\pi^2} - 1)^{-\frac{1}{2}} - 1$. It is easy to see that $sp^2 < sp^1$ for $\pi \in [0; 1]$. In other words, even for bonds issued by the same bank, the absolute value of the spread may vary across maturities. This time effect is linked to the risk-aversion of the investors,¹⁵ but it does not reflect different probabilities of bankruptcy of the bank, in which I am actually interested. In addition, since it is more difficult to assess the creditworthiness of a bank in the distant future, risk-averse investors may require an additional risk premium for longer-term bonds. In any case, it is important to control for the maturity effect in the estimations.

Finally, I include the coupon ($coupon_i$) of the bond as explanatory variable. In theory, the coupon should affect spreads in two contrasting ways. On the one hand bonds with high coupons are considered to be less risky (lower duration and volatility) than low coupon bonds. On the other hand, in some countries zero bonds may have tax advantages, since capital gains are tax-free while coupons are treated as taxable income.¹⁶

The Swiss banking sector and its regulatory background

The very particular structure of the Swiss banking sector could play a major role in this context, and therefore needs to be analyzed further. First, the market is highly concentrated. In 2004, the two big banks, UBS and CSG, owned about 75% of the total assets in the sector. Because of their systemic relevance those institutions may be thought to be “too-big-to-fail” (TBTF). If investors speculate that a TBTF bank would always be bailed out by the authorities, there is no reason to require a risk premium for lending money to that bank and as a result market discipline would be ineffective. Rime (2005), for instance, finds that the (presumed) TBTF status of a bank has a positive impact on issuer ratings. However, by looking at the size of UBS and CSG – total assets being a multiple of Swiss GDP – one may wonder if those banks are also “too-big-to-be-rescued”, which would partially offset the distorting effect of the TBTF problem. To control for this effect I include the log value of total assets ($ta_{b,t}$) as explanatory variable in the model specification.

Second, many of the bonds included in the sample were issued by cantonal banks. Those institutions are protected by an implicit (Banque Cantonale Vaudoise, BCV), limited (Banque Cantonale de Genève, BCGe) or explicit state guarantee. This may also limit the effectiveness of market discipline, since debt investors are again not at risk. As a corollary, they also have no incentives to monitor those banks. Sironi (2003) provides empirical evidence that state guarantees distort the

¹⁵With risk-neutrality $sp^2 = sp^1$.

¹⁶In Switzerland capital gains on tradable debt are considered as taxable income if the coupon yield is low when compared to the total return. Moreover, coupons are subject to withholding tax.

2 The prospects of market discipline for banks. Evidence from the Swiss bond market

information content of market data. In particular, he finds that, at issuance, subordinated debt spreads are about 40 basis points lower for European public-sector banks.

Finally, as in almost every country, depositors in Switzerland have the highest seniority status in the case of default and are partially protected by deposit insurance.¹⁷ While this protection reduces the incentives for depositors to monitor banks, the priority rule – if respected – enhances the disciplining role of the debenture market, since all bondholders groups are subordinated (albeit to a different degree) and therefore more exposed to losses. This effect should be captured by $deposit_{b,t}$, the total amount of deposits in relation to total assets.

2.2.2 The empirical model

After having outlined why some exogenous variables should have an effect on bond spreads, I need to be more specific on the functional form of the relationship.

The following two equations are the starting point of my empirical analysis

$$risk_{b,t} = g(X_{b,t}, \alpha, \varepsilon_{b,t}) \quad (2.9)$$

$$sp_{i,b,t} = h(risk_{b,t}, Z_{i,t}, \beta, u_{i,b,t}) \quad (2.10)$$

where $risk_{b,t}$ represents the risk of bank b at time t and $sp_{i,b,t}$ is the spread of bank b 's debenture i .¹⁸ $X_{b,t}$ and $Z_{i,t}$ are vectors of explanatory variables, α and β vectors of parameters. The error terms, $\varepsilon_{b,t}$ and $u_{i,b,t}$, are assumed to be serially uncorrelated.

In principle, bank supervisors and investors would like to estimate an equation like (2.9), the relationship between the risk of a bank and various explanatory variables. Unfortunately, this is a rather difficult task. A lot of the risk determinants such as the risk appetite of the bank's managers, cannot be perfectly observed even by the supervisor. In addition, the functional form of $g(\cdot)$ is a priori unknown and I do not know any model that can help solve this problem in a satisfactory manner. Finally, it is difficult to calibrate such a model, as the true value of $risk_{b,t}$ is hardly measurable and observable even ex post. Regulators therefore look for indicators as proxies for risk. Such indicators include bond spreads. In the model developed in Section 2.1, I showed how spread and risk are related, namely

$$risk_{b,t} = g'(sp_{i,b,t}, X'_{i,t}, \alpha', \varepsilon'_{b,t}) \quad (2.11)$$

However, I still lack a reliable measure of $risk_{b,t}$. Balance sheet ratios could be used as proxies for risk, but using them as a dependent variable entails a considerable danger of misspecification of the model, since accounting figures depend on factors other than risk. One way to solve this problem is to reverse this relationship. As economic theory can better explain the determinants of

¹⁷See Birchler and Maechler (2002) for an exhaustive description of the Swiss depositor protection system.

¹⁸Above I defined $sp_{i,b,t}$ as $sp_{i,t}$, i.e., without the bank-specific subscript. In this section it is important to make this distinction because of the special form of the variance covariance matrix (see below).

bond spreads, estimating a model like (2.10) should lead to more accurate results. But what is the link between equation (2.10) and (2.11)? In my model $\tilde{X}_{i,t} = Z_{i,t}$ and $\tilde{g}()$ is the inverse of $h()$, but in reality things are more complex. In Section 2.1, for instance, I considered only the relationship between bond spreads and the endogenous risk of a bank, but it is well known that the spread is determined by (at least) three sources of risk, namely systematic, default and liquidity risk. Moreover, the exact functional form of $h()$ is unknown. Although my model is undoubtedly too simple to be used directly in specifying (2.10), one can obtain some important insights by looking at equations (2.5) and (2.6). For example, even in this very simple model the relationship between spreads and bank risk is not linear. This aspect was already pointed out by Black and Cox (1976) and emphasized later by Gorton and Santomero (1990). Black and Cox showed in particular that the price of subordinated debt is initially a convex function of the firm's value but if this value is high enough the dependence becomes concave. Given that nonlinearities play an important role but I do not know the functional form of $h()$, nonparametric estimation procedures could represent a useful way to solve the problem. Unfortunately the sample is rather small so that applying those econometric techniques could be inefficient and lead to imprecise results. Therefore, I have to rely upon more traditional approaches for calculating β . That is, I will estimate

$$sp_{i,b,t} = h(R_{b,t}, Z_{i,t}, \beta, u_{i,b,t}) \quad (2.12)$$

by making simplifying assumptions about the functional form of $h()$. $R_{b,t}$ is a vector of the instruments proxying $risk_{b,t}$ that were described in Section 2.2.1.

On top of this, there are two additional technical issues to control for. First, the covariance matrix of the error terms Ω is likely to have a non-standard form. As some bonds and some banks are traded more than others and my measures of the liquidity premium are possibly not perfect, the errors are likely to be heteroskedastic. Also particular in this setting is the correlation across panels (bonds) due to the fact that several bonds in the sample were issued by the same bank. In other words, suppose that

$$u_{i,b,t} = \tilde{v}_b + \hat{v}_{i,b} + v_{i,b,t}$$

where \tilde{v}_b and $\hat{v}_{i,b}$ are, respectively, the bank-specific and the bond-specific part of the error term. If I assume that \tilde{v}_b and $\hat{v}_{i,b}$ do not depend on t and that each component is independent, I have

$$\begin{aligned} \tilde{v}_b &\sim iid[\tilde{\eta}_b, \tilde{\sigma}_b^2] \text{ for } b = 1, \dots, B \\ \hat{v}_{i,b} &\sim iid[\hat{\eta}_{i,b}, \hat{\sigma}_{i,b}^2] \text{ for } i = 1, \dots, I_b \text{ and } b = 1, \dots, B \\ v_{i,b,t} &\sim iid[0, \sigma^2] \text{ for all } b, i \text{ and } t. \end{aligned}$$

where I_b is the set of bonds issued by bank b . To see the implications, suppose that both $\tilde{\eta}_b$ and $\hat{\eta}_{i,b}$ equal zero. Then $E(u_{i,b,t}u_{i,b,t}) = \tilde{\sigma}_b^2 + \hat{\sigma}_{i,b}^2 + \sigma^2$, which depends on both b and i . In addition, $E(u_{i,b,t}u_{j,b,\tau}) = \tilde{\sigma}_b^2$ for all $i \neq j$ and $t \neq \tau$. That is, errors would be heteroscedastic and the panels

correlated. This non-standard form of Ω is shown in Appendix 2.5.5. I tested the assumptions that the errors are homoscedastic and uncorrelated across panels with a LR test, but I had to reject the null. Because of the special form of the covariance matrix, I cannot use standard procedures like the random effects model to estimate Ω . Instead, I will rely on correlated panels corrected standard errors (PCSEs).

The second issue is related to the fact that some of the explanatory variables may be endogenous. One potential problem lies in the fact that my measures of bank risk are not optimal. The variables observed may not be the variables to which economic agents react, but some proxies of them. If this is true, the unobservable term contains a measurement error which can be correlated with the regressors. Moreover, although I try to include all determinants of bond spreads in my econometric model, unobserved heterogeneity across banks ($\tilde{\eta}_b \neq 0$) and bonds ($\hat{\eta}_{i,b} \neq 0$) may still bias the analysis. If measurement errors and unobserved heterogeneity are constant over time, I can eliminate both problems by using a fixed effects approach. The endogeneity bias happens to be even more serious to deal with if $E(v_{i,b,t} | R_{b,\tau}, Z_{i,\tau}, \tilde{v}_b, \hat{v}_{i,b}) \neq 0$ for some τ . Especially for $\tau = t$ the exogeneity of some regressors may be questionable. For instance, the liability structure of the balance sheet may depend on the current perspectives of the bank in the capital markets, which are reflected in the spread. In those cases one could use instruments – e.g., lagged values – for the problematic explanatory variables. On the other hand, in efficient markets financial prices such as bond spreads are driven exclusively by investors' expectations.¹⁹ In addition, for indirect market discipline it is important to know what spreads can tell supervisors about the future. Since expectations are not observable, a solution would be to proxy them with present and leading observations. So there is a trade-off between the choice of a good econometric versus a good economic model. I assume that expectations play a major role with regard to the asset risk and performance variables as well as the general economic environment. Therefore, I use leads only for those variable classes. In fact, the liability structure of a bank remains relatively constant over time and changes in systematic risk factors are extremely difficult to predict, so present figures should already summarize all available information about the future. More importantly, for those two variable classes the endogeneity problem may be particularly relevant.

A final note on multicollinearity. Since some regressors – such as the ratio of provisions and non-performing loans to total assets – are highly correlated, I use absolute differences instead of leads and lags as explanatory variables. Secondly, the economic cycle regressors – such as the overall interest rate margin or the unemployment rate – do not vary across groups. In order to reduce the multicollinearity problem, for those variables I include only leads in the estimation.

¹⁹The importance of testing whether market prices reflect ex ante risks was already emphasized by Morgan and Stiroh (2001).

2.2.3 Data set and sample characteristics

I run all the estimations with STATATM. My sample includes 236 senior and 53 subordinated non-callable, non-puttable and non-convertible bonds of 30 banks. In contrast to most studies on indirect discipline in the debt market, I include subordinated as well as senior bonds in the empirical analysis. In this way I hope to enhance the randomness of the sample, as about two third of the subordinated debentures were issued by the two big banks. On the other hand, the model in Section 2.1 made clear that the functional form of the relationship between spread and bank risk differs according to the seniority status of debt. Hence, using the same empirical model for both senior and subordinated debt may involve some danger of misspecification. However, two points are worth noting. First, under Swiss legislation traded bonds are subordinated to a large fraction of the deposits. Hence, both types of bonds can be viewed as subordinated debt. Second, priority rules – with the exception of deposits – are often violated in reality.²⁰

Table 2.1: Summary statistics

Spreads in basis points; all ratios in percent.				
Variable	Mean	Std. Dev.	Min.	Max.
$sp_{i,t}$	48.13	44.65	-164.06	464.34
$roa_{b,t}$	0.33	0.57	-3.54	4.12
$provision_{b,t}$	3.62	2.25	0.48	11.95
$newprov_{b,t}$	0.66	0.93	-0.06	9.34
$npl_{b,t}$	3.29	3.12	0	17.58
$ta_{b,t}$	10.51	1.67	7.04	14.13
$busrisk_{b,t}$	53.40	12.13	17.67	100.31
$liquidity_{b,t}$	62.28	70.30	-13.14	459.28
$ebk_{b,t}$	0.11	0.31	0	1
$deposits_{b,t}$	29.43	14.96	0.99	66.14
$senior_{b,t}$	15.98	9.76	0	34.87
$junior_{b,t}$	0.88	1.03	0	4.17
$amount_i$	162.70	105.43	40	1000
$coupon_i$	4.13	1.01	1.75	7.75
sub_i	0.21	0.41	0	1

All the selected securities are denominated in Swiss francs and were issued between 1990 and 2004. Hence, I do not have to deal explicitly with exchange rate risk. Since I lack a comprehensive bank balance sheet database before 1995, I restrict my analysis to spreads calculated between 1995 and 2006. Accordingly my unbalanced panel consists of 21,975 monthly spread observations. In the estimations I need to restrict my sample even further, as some of the balance sheet variables are available only on a quarterly or yearly basis. Income statement figures and data on non-performing loans, e.g., are collected once a year. On the other side, the time dimension plays an important role in my analysis. As a trade-off I estimate the model with quarterly data. For yearly series I replaced the three missing data with the last available observation.

²⁰See Longhofer and Carlstrom (1995) for an overview of the US case.

2 The prospects of market discipline for banks. Evidence from the Swiss bond market

Table 2.1 reports the summary statistics of the key variables in my sample. All ratios are in percent. The majority of the spread observations comes from big banks (20%) and cantonal banks (63%). As one would expect, subordinated debentures have a higher average spread (83bp) than senior bonds (39bp). The low spread averages are possibly due to the high number of bonds issued by cantonal banks, many of which enjoy an explicit state guarantee. In fact, the mean spread of senior debt increases to 58bp if I exclude fully guaranteed cantonal banks from the sample. The state guarantee and liquidity premia may also explain the occurrence of negative spreads in the data set. The repayment period varies between 3 and 15 years, averaging roughly 9 years. The mean maturity is higher for subordinated debentures, which is probably due to the Basel Capital Accord rules, which require a minimum original term to maturity of over five years for subordinated instruments in order to qualify as Tier 2 capital.

The bank sample is highly heterogeneous. Since I include major global players as well as relatively small regional credit banks and private institutions in the estimations, big differences in the dimensions and structure of the balance sheet emerge. For example, at the end of 2002 the totals assets of the smallest bank of the sample, Banque Cantonale du Jura (BCJ), amounted to about 2 billion Swiss francs, while for UBS the balance sheet total exceeded 1,000 billion. Accordingly, the amount of the issue ranges between 40 (BCJ) and 1,000 (CSG) million francs.

2.3 Results

2.3.1 Linear specification

It is often argued that reduced-form models do a better job in pricing credit risk than structural models.²¹ Duffie and Singleton (2003) show that under some assumptions – in particular when the recovery rate is zero – reduced-form models imply that the yield spread of a defaultable bond equals the hazard rate underlying the default process. By further modeling the hazard rate as an affine function of the exogenous variables described above, one could then use a linear specification to estimate (2.10). Accordingly, as a first approximation, in this section I assume that $h()$ is linear. In order to control for unobserved heterogeneity at the bank level ($\tilde{\eta}_b \neq 0$) I introduce a dummy variable for each bank. The results of those linear regressions are shown in Table 2.2. I calibrated the model for two different samples: (A) banks that do not enjoy an explicit full state guarantee and (B) fully insured cantonal banks.²² The rationale behind this subdivision is that bond spreads of cantonal banks may be biased by the state guarantee and therefore do not depend on default risk. Hence, running the model including all banks would increase the danger of model misspecification. Two cantonal banks, BCV and BCGe, are included in sample (A). BCV does not enjoy an explicit state guarantee. Therefore, it can be considered as a “normal” bank. In the case of BCGe, the

²¹See Section 2.1.

²²For reasons of space, only the results for the variables that concern the main hypothesis are shown. The whole set of results can be found in Appendix 2.5.6.

state guarantee applies only to deposits and consequently should not affect the risk perception of bondholders.²³

Let us first take “normal” banks. The results for the performance variables proxying credit risk are mainly in line with the expectations. The coefficient of $newprov_t$ and Δnpl_{t+4} are significant at the 1% level and have the predicted sign. If the ratio of a bank’s non-performing loans to its total lendings increases by five percentage points over the next year, for instance, the spread of bonds issued by that bank will go up by roughly 60 bp.²⁴ In other words, investors are forward-looking, update their assessment of the bank’s credit-risk profile on the basis of those variables and adjust the risk premia on bond prices accordingly. The current absolute amount of provisions to total credits is also negatively correlated with bond prices, albeit at a 10% significance level. A priori somewhat surprising is the positive coefficient on current and future profitability. However, as pointed out in Section 2.2, high profits are not necessarily good news for bondholders as they may reflect excessive risk-taking.

Let me now consider the asset risk variables. Not surprisingly the coefficient on ta_{t-1} is insignificant. Since I included a dummy for each bank, possible size effects are mainly captured by those dummies. Hence, I cannot reject the hypothesis that a TBTF effect exists. In the next section I analyze this issue in greater detail. The positive and highly significant coefficient on $busrisk_t$ indicates that the current asset mix of a bank is also a relevant factor affecting the spread. My measure of bank liquidity is positively correlated with bond spreads as well, albeit not significantly. I will show below that this result is mainly due to the behavior of the two big banks, which accumulate liquidity during periods of stress.

While the current SFBC rating is reflected in the spread, the coefficient Δebk_{t+1} is not significantly different from zero. This is an important result. Investors detect and incorporate in their solvency assessment of a bank information which is not directly related with the standard balance sheet variables included in the estimation. However, they do not appear to be able to detect those problem banks before supervisors do. Still, one has to be careful with this conclusion. Δebk_{t+1} may be correlated with other variables (Δnpl_{t+4} for instance) and hence the effect on the spread may be captured by the latter. More importantly, the SFBC may use bond spreads or other market indicators as information sources or as a trigger for special audits or management reports. It thus seems reasonable to expect that, if the bond spreads of a bank increase and the SFBC detects that a significant problem really exists, the bank will soon be put on the “black list”.

The liability structure of the balance sheet is also a relatively important determinant of bond prices. The higher the amount of (senior) debt outstanding, the higher the risk premium required

²³Clearly, one can argue that even a limited or implicit state guarantee *does* affect the risk perception of investors, as they may expect that the canton will rescue a troubled state bank anyway. Indeed, the Canton of Vaud virtually bailed out BCV in 2002. However, it is reasonable to assume that the state guarantee effect is strongest when investors are protected by law.

²⁴The term “expectations” is somewhat misleading, especially for those variables measured on a yearly basis. In fact investors can update their information about net income or non-performing loans from the quarterly or half-year reports of each bank, when available.

Table 2.2: Base-model estimation results

Bank bond spreads are regressed on a set of explanatory variables using a PCSE model with bank-specific unobserved heterogeneity. All ratios are in percent.				
Variable	Sample:	(A) "Normal" banks	(B) Insured banks	
		Coefficient (Std. Err.)	Coefficient (Std. Err.)	
<i>Performance:</i>				
roa _t		15.79** (3.90)	19.17 (13.31)	
Δroa _{t+4}		5.08 [†] (2.76)	-1.34 (14.05)	
provision _t		3.13 [†] (1.72)	10.56** (2.71)	
Δprovision _{t+4}		2.63 (2.35)	4.81 [†] (2.54)	
newprov _t		13.21** (2.33)	-5.52 (6.31)	
npl _t		-1.19 (1.69)	-0.86 (1.27)	
Δnpl _{t+4}		11.92** (2.62)	-3.52** (1.10)	
<i>Assets:</i>				
ta _{t-1}		-3.48 (8.30)	29.46 (34.98)	
busrisk _t		1.05** (0.29)	-0.15 (0.37)	
Δbusrisk _{t+4}		0.34 (0.33)	0.59 [†] (0.32)	
liquidity _{t-1}		0.04 (0.03)	-0.01 (0.01)	
<i>SFBC rating:</i>				
ebk _t		18.50** (7.09)	1.04 (9.74)	
Δebk _{t+1}		15.68 (9.64)	27.98 [†] (14.55)	
<i>Liability structure:</i>				
deposit _{t-1}		0.19 (0.74)	1.33 (0.86)	
seniord _{t-1}		2.03* (0.94)	-0.62 [†] (0.36)	
juniord _{t-1}		10.42 (6.74)	-11.15 (7.51)	
Dseniord _{t-1}		1.01** (0.26)	-0.20 (0.43)	
Djuniord _{t-1}		-2.34 (2.32)	8.12 (7.82)	
Number of obs.		2266	2841	
Number of banks		13	17	
R ²		0.74	0.45	
χ ²		17672.0	3089.2	
Significance levels : [†] =10% *=5% **=1%				

by debt-holders. In line with every structural model based on Merton's approach, a high leverage means that less capital is available as a security buffer to cover possible losses. In addition, the option part of equity value becomes relatively more important, which induces the bank to take more risks. The coefficients on $Dseniord_{t-1}$ and $Djuniord_{t-1}$ also have the predicted sign, although the latter is not significantly different from zero.

With respect to the variables not directly related to idiosyncratic default risk, the results confirm that the economic cycle and systemic risk factors are important determinants of bond spreads. However, the interest rate and the slope of the yield curve have a positive coefficient, which is surprising since, along with the Merton model and empirical evidence²⁵, changes in credit spreads and in the yield curve should be negatively correlated. The reason for this apparent paradox lies in the fact that, in order to enhance exogeneity, I used the lags of those variables. If I had used current values, both coefficients would have become negative. This effect is possibly due to the lower liquidity of corporates compared to treasuries, whose prices react more quickly to shocks. When bank bond yields adjust only after a lag to their equilibrium values following a change in the term structure, the effect of the risk-free rate on the spread would exhibit such a pattern. As expected, subordinated debentures trade with a price discount over senior bonds. At first sight, the average premium of 24bp seems to be quite low. Since the difference in yield between subordinated and senior debt tends to be positively correlated with credit quality, this result may be due to the high average credit quality of Swiss banks. Finally, my model is not very accurate in explaining bond liquidity risk, as no specific proxy is statistically significant at at least the 5% level.

On the whole, the findings support the view that bond prices are risk-sensitive, i.e., the market monitoring hypothesis cannot be rejected. On the other hand, there appears to be no scope for indirect market discipline, as supervisors cannot learn very much from the market. Because of the endogeneity problem described above, however, this conclusion should be treated with caution.

Now let us consider fully insured cantonal banks. The first thing that leaps out is the difference in the R-squared statistics between the two samples. While my model turns out to be relatively good in describing spreads of "normal" banks, its explanatory power shrinks considerably when only fully insured state-owned banks are examined. Moreover, one of the two variables of interest which are statistically significant at least at the 5% level, Δnpl_{t+4} , has the wrong sign. Interesting are the slightly significant coefficients on $\Delta provision_{t+4}$, $\Delta busrisk_{t+4}$ and Δebk_{t+1} , which seems to suggest that cantonal banks' investors too are somewhat forward-looking. In general, however, the return on debt issued by state-owned institutions appears to depend relatively more on systematic and liquidity risk factors. The low R-squared is probably due to the fact that my measures of bond liquidity are quite imprecise and fail to accurately control for the liquidity premium.

I interpret these results as evidence for the distorting effects of the state guarantee from the perspective of market discipline. Some caution is needed, though. My cantonal bank sample is rather uninformative, as no institution other than BGV and BCGe had substantial problems during the period under analysis. This may explain the low volatility of cantonal banks' bond prices: The

²⁵See Van Landschoot (2004) for an overview.

average standard deviation of senior bond spreads equals 0.40 in sample (A) and 0.28 in sample (B). Yet, it is difficult to say whether this low volatility is due to the state guarantee or the absence of major difficulties at cantonal banks. In order to derive a definitive conclusion on the effect of the state guarantee, one has to “hope” that some public bank will encounter difficulties and see how investors react. On the other hand, it is worth pointing out that bonds issued by BCV and BCGe traded at a statistically significant price discount of about 20bp even before the difficulties at these two institutions became public. To sum up, while the state guarantee appears to limit the flow of risk-related information in financial prices, one cannot rule out the possibility that cantonal banks’ investors may not be concerned with fundamentals at all.

2.3.2 Discussion and robustness tests

The linear model estimated in the previous subsection did not control for several factors, such as bond-specific unobservable effects that, in principle, could have biased the analysis. In addition, my bank sample was rather small and very heterogeneous. Therefore I performed various checks to assess the robustness of the main results.

I first estimated the model using a sample of parent banks that do not enjoy an explicit state guarantee. The reason is that spreads of subsidiaries should a priori depend more on the solvency of the parent company than on their own risk profile. The other side of the coin is that I am then left with only 8 banks and about half of the observations in the sample stem from big banks. As shown in Table 2.3 below, the results obtained with this smaller data set are very similar to previous ones, although the t-values are often lower. This is mainly because the estimates are less precise (high standard deviations), which is possibly due to the limited randomness of my bank sample. More difficult to justify is the significant coefficient on ta_{t-1} . One reason may be that investors require an additional risk premium for banks that expand rapidly.

As a further test, I excluded the two big banks from the base sample of “normal” banks, in order to focus more on relatively small institutions. The estimation results are shown in the middle columns of Table 2.3. Again, the findings are very similar, although some aspects are worth pointing out. First, Δroa_{t+4} is no longer significant. One reason for this result may be that the two big banks publish detailed quarterly reports. Investors therefore have the opportunity to regularly update their risk assessments and profit expectations. Conversely, smaller Swiss banks only issue annual and, in some cases, half-year reports. The latter are usually quite rudimentary and therefore not very informative. On the other hand, $provision_t$ becomes significant. Next, note the negative – albeit insignificant – coefficient on the liquidity variable. One can argue that only smaller banks need a buffer to absorb liquidity shocks, while big banks may have external sources of funds – such as the lender of last resort²⁶ – which are not always available to other institutions. Moreover, big banks tend to accumulate liquidity during periods of stress, which explains the positive coefficient

²⁶A bank may benefit from the Swiss lender of last resort – the Swiss National Bank – only if it is solvent and systemically relevant. Clearly, UBS and CSG are more likely to meet this last criterion than other banks.

on $liquidity_{t-1}$ in the parent banks sample. Finally, the asset mix no longer seems to play a role. The fact that the activity of most small banks is limited to a specific region and business area (off-shore private banking or mortgage business in Switzerland) and hence $busrisk_t$ does not vary a lot within a panel may explain this result.

Finally, I excluded from sample (A) the banks that are essentially involved only in asset management. Indeed, some of the explanatory variables such as provisions and NPL are key indicators of credit risk, and the presence of such banks in the sample may, in principle, reduce the explanatory power of the estimation. As shown in the last columns of Table 2.3, though, excluding asset managers from the base sample does not alter the main findings. An exception is $busrisk_t$, which becomes insignificant.

Next, I controlled for the importance of unobserved heterogeneity at the bond level by introducing a bond-specific dummy. The results are very similar to those of the basic model (see Table 2.4). In other words, constant bond-specific variables other than the amount issued and the coupon do not seem to be an important determinant of the spreads. Hence, a substantial part of the market liquidity premium appears to be time-varying and thus it is extremely difficult to measure this premium accurately.

As a benchmark, I also estimated the model without any dummy. The hardly justifiable negative coefficient on $busrisk_t$ suggests that some bank-specific variables are correlated with unobserved heterogeneity at the bank level. Thus, the results of the model with no bank-specific dummy may be biased. This problem also remains if I introduce a big bank dummy. Consequently, one has to be careful in interpreting the negative coefficient on ta_{t-1} as evidence that the TBTF issue may influence investors' risk perception for big institutions. Even if that coefficient were unbiased, debt investors could simply consider large banks safer for a number of other reasons (e.g., better supervision, broader diversification or more sophisticated risk management). Moreover, bonds of bigger institutions may exhibit a lower market liquidity premium. For sure, debt securities of the two big banks cannot be considered risk-free. In October 2002, investors required a return premium for subordinated bonds issued by CSG which was almost 150 basis points above that of UBS (see Figure 2.1). If investors have no doubt that both institutions are TBTF, such price discounts were not justifiable. On the other hand, on the basis of my results I cannot reject the hypothesis that a TBTF distortion exists.

Finally, I respecified the basic model without using leads as explanatory variables. As emphasized in Section 2.2.2, leads may be endogenous to the spread, which would bias the estimation. As shown in Table 2.5, however, the results are very similar to those of the original model, especially for sample (A). Hence, I can conclude that using leads as explanatory variables does not materially bias the estimations. With respect to the sample of insured banks more differences between the original and the respecified model arise. This is not surprising and may be interpreted as further evidence that, because of the state guarantee, my econometric model is not very suitable for explaining the variation in the spreads of bonds issued by insured banks.

Table 2.3: Panel sensitivity

Bank bond spreads are regressed on a set of explanatory variables using a PCSE model with bank-specific unobserved heterogeneity. All ratios are in percent.				
Variable	Sample:	(C) Parent banks	(D) Small banks	(E) Credit banks
		Coefficient (Std. Err.)	Coefficient (Std. Err.)	Coefficient (Std. Err.)
<i>Performance:</i>				
roa _t		19.54** (5.74)	8.14† (4.21)	31.44** (5.62)
Δroa _{t+4}		6.76† (4.10)	1.60 (2.74)	10.99** (3.32)
provision _t		3.37 (2.07)	4.46* (2.25)	1.95 (1.85)
Δprovision _{t+4}		3.53 (3.40)	1.51 (2.69)	2.55 (2.93)
newprov _t		12.20** (2.65)	8.59** (2.51)	21.49** (3.66)
npl _t		2.92 (3.55)	-0.34 (1.56)	0.99 (2.50)
Δnpl _{t+4}		13.64** (3.32)	16.18** (3.02)	13.41** (2.96)
<i>Assets:</i>				
ta _{t-1}		29.24* (12.74)	-3.68 (7.09)	7.06 (8.95)
busrisk _t		1.25** (0.44)	0.82 (0.56)	0.03 (0.65)
Δbusrisk _{t+4}		0.33 (0.39)	0.44 (0.41)	-0.37 (0.69)
liquidity _{t-1}		0.07* (0.03)	-0.01 (0.04)	0.07† (0.04)
<i>SFBC rating:</i>				
ebk _t		17.13* (8.52)	21.68** (7.48)	15.22† (8.69)
Δebk _{t+1}		10.96 (11.43)	14.58 (9.18)	17.76† (10.1)
<i>Liability structure:</i>				
deposit _{t-1}		-0.77 (1.12)	1.24 (0.89)	0.53 (0.84)
seniord _{t-1}		6.83** (1.84)	1.86† (1.02)	3.09** (1.01)
junior _{t-1}		21.15* (10.31)	41.81** (13.07)	14.33† (7.95)
Dseniord _{t-1}		1.12** (0.26)	-0.66 (0.86)	1.08** (0.27)
Djunior _{t-1}		-1.43 (2.59)	-26.52* (10.58)	-4.57† (2.51)
Number of obs.	1952	1426	2067	
Number of banks	8	11	9	
R ²	0.76	0.78	0.76	
χ ²	1531.7	4576.1	1485.9	
Significance levels : †=10% *=5% **=1%				

Table 2.4: Unobserved heterogeneity

Bank bond spreads are regressed on a set of explanatory variables using a PCSE model with different assumptions about unobserved heterogeneity. (a) unobserved heterogeneity at the bond level; (b) no unobserved heterogeneity; (c) unobserved heterogeneity relevant with respect to big bank group. The sample consists of “normal” banks. All ratios are in percent.						
Variable	(a) Bond specific		(b) No dummy		(c) BB-dummy	
	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)
<i>Performance:</i>						
roa _t	16.30**	(3.70)	1.48	(2.65)	2.19	(2.53)
Δroa _{t+4}	4.40 [†]	(2.65)	-2.31	(2.63)	-2.08	(2.58)
provision _t	4.69**	(1.75)	2.80	(1.81)	1.86	(2.14)
Δprovision _{t+4}	2.87	(2.35)	0.98	(2.57)	0.90	(2.52)
newprov _t	12.74**	(2.29)	6.94**	(1.54)	7.94**	(1.64)
npl _t	-1.59	(1.65)	0.79	(1.14)	0.85	(1.42)
Δnpl _{t+4}	11.33**	(2.55)	12.31**	(2.85)	12.23**	(2.83)
<i>Assets:</i>						
ta _{t-1}	-0.78	(7.73)	-12.17**	(1.59)	-13.69**	(2.69)
ta _{t-1} if big bank	—	—	—	—	-5.50**	(1.72)
busrisk _t	1.24**	(0.33)	-0.77*	(0.19)	-1.34**	(0.25)
busrisk _t if big bank	—	—	—	—	1.60*	(0.63)
Δbusrisk _{t+4}	0.46	(0.33)	-0.19	(0.37)	-0.48	(0.38)
liquidity _{t-1}	0.03	(0.03)	0.04 [†]	(0.03)	-0.12**	(0.04)
liquidity _{t-1} if big bank	—	—	—	—	0.20**	(0.05)
<i>SFBC rating:</i>						
ebk _t	17.01*	(7.13)	9.29 [†]	(5.63)	-2.15	(7.49)
Δebk _{t+1}	14.89	(9.81)	11.08	(9.83)	7.31	(10.31)
<i>Liability structure:</i>						
deposit _{t-1}	-0.88	(0.69)	0.34**	(0.11)	0.15	(0.14)
seniord _{t-1}	1.33 [†]	(0.80)	0.21	(0.19)	0.74*	(0.29)
juniord _{t-1}	18.15*	(8.38)	17.19**	(3.74)	18.12**	(3.74)
Dseniord _{t-1}	3.64*	(1.76)	0.90**	(0.29)	0.88**	(0.28)
Djuniord _{t-1}	-8.64	(6.78)	-5.24*	(2.20)	-5.45*	(2.16)
<hr/>						
Number of obs.	2266		2266		2266	
Number of banks	13		13		13	
R ²	0.77		0.72		0.73	
χ ²	5215.7		807.5		965.9	
<hr/>						
Significance levels :	†=10%	*=5%	**=1%			

Table 2.5: Base-model with no leads

Bank bond spreads are regressed on a set of explanatory variables using a PCSE model with bank-specific unobserved heterogeneity. All ratios are in percent.				
Variable	Sample:	(A) "Normal" banks	(B) Insured banks	
		Coefficient (Std. Err.)	Coefficient (Std. Err.)	
<i>Performance:</i>				
roa _t		7.67 [†] (4.09)	-1.31 (16.85)	
Δroa _{t-4}		-1.04 (2.51)	24.17 (17.25)	
provision _t		4.25** (1.51)	7.18* (3.33)	
Δprovision _{t-4}		2.63 (2.36)	-4.23 [†] (2.32)	
newprov _t		10.28** (2.65)	-6.77 (7.27)	
npl _t		-2.43 (1.75)	2.77* (1.34)	
Δnpl _{t-4}		11.38** (2.96)	1.54 (1.38)	
<i>Assets:</i>				
ta _{t-1}		9.77 (8.00)	29.36 (31.88)	
busrisk _t		1.39** (0.46)	-0.35 (0.38)	
Δbusrisk _{t-4}		-0.31 (0.48)	-0.15 (0.34)	
liquidity _{t-1}		-0.03 (0.03)	-0.05** (0.02)	
<i>SFBC rating:</i>				
ebk _t		13.31 [†] (7.79)	-11.60 (9.30)	
Δebk _{t-1}		-4.72 (10.81)	12.35 (15.22)	
<i>Liability structure:</i>				
deposit _{t-1}		-3.53** (0.92)	1.50 [†] (0.81)	
seniord _{t-1}		-0.35 (0.95)	0.23 (0.37)	
juniord _{t-1}		12.63* (6.37)	8.43 (10.78)	
Dseniord _{t-1}		1.10** (0.31)	1.02* (0.42)	
Djuniord _{t-1}		-5.87* (2.76)	-7.52 (10.84)	
Number of obs.		1942	2581	
Number of banks		13	17	
R ²		0.74	0.36	
χ ²		5973.6	3390.1	
Significance levels : †=10% *=5% **=1%				

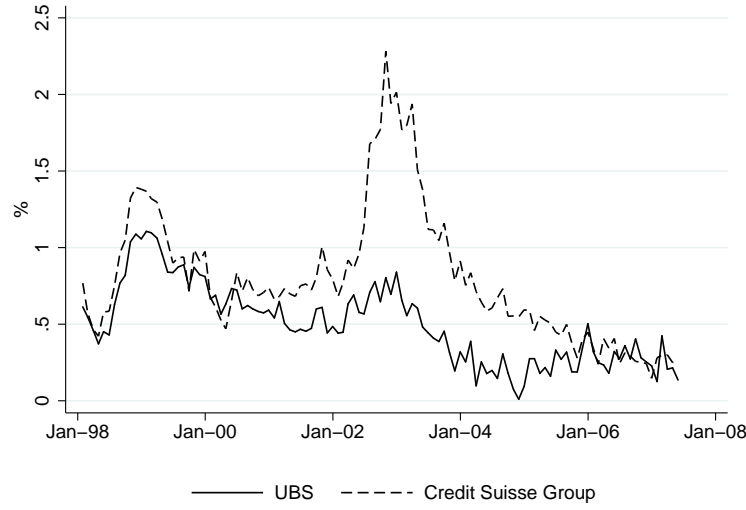


Figure 2.1: Average subordinated bond spreads

2.3.3 Nonlinearities

Even in my simple theoretical model, the relationship between bank risk and bond spread is not linear. Thus, a natural question arises: Does the linear specification produce unbiased results? The problem is that in reality I do not know the exact functional form of $h(\cdot)$. Since the sample is small and the number of independent variables relatively large, nonparametric estimation procedures are likely to be rather imprecise. However, in order to get an idea of the importance of the nonlinear components of $h(\cdot)$, I estimated (2.12) using a Generalized Additive Model (GAM) of the form

$$sp_{i,b,t} = h_0 + \sum_j h_j \left(x_{i,b,t}^j \right) + u_{i,b,t}$$

where $x_{i,b,t}^j \in X_{i,b,t} = R_{b,t} \cup Z_{i,t}$. $h_j(\cdot)$ is a non-parametric function estimated via a cubic smoothing spline. This GAM model produced the results shown in Table 2.6.

According to the GAM model, the nonlinear component is indeed important: For the majority of the regressors, the difference in normalized deviance between the GAM and the linear model (the “gain”) is significantly different from zero. This is true especially for the performance and the liability structure variables. To get an idea of the shape of the nonlinear relationship between the spreads and the regressors, I estimated the basic model using linear splines. Linear splines allow us to shape the relationship between y and x as a piecewise linear function. To keep things simple, I decomposed the variables in three segments. In the first specification (Table 2.7) the knots are placed at a priori “strategic” relevant points in the whole sample distribution of each variable, also making sure that each segment contains enough observations. For example, for roa_t the knots are placed at 0% (profit versus losses) and at 0.5%. In the second specification (Table 2.8) the knots

Table 2.6: Importance of nonlinearities

Comparison between the GAM and the linear model with bank-specific unobserved heterogeneity.		
Variable	Gain [†]	Significance [‡]
<i>Performance:</i>		
roa _t	59.64	1%
Δroa _{t+4}	17.63	1%
provision _t	13.10	1%
Δprovision _{t+4}	4.83	10%
newprov _t	72.84	1%
npl _t	38.54	1%
Δnpl _{t+4}	14.17	1%
<i>Assets:</i>		
ta _{t-1}	2.98	not significant
busrisk _t	4.87	10%
Δbusrisk _{t+4}	7.79	5%
liquidity _{t-1}	7.53	5%
<i>Liability structure:</i>		
deposit _{t-1}	143.90	1%
seniord _{t-1}	28.08	1%
juniord _{t-1}	16.99	1%
Dseniord _{t-1}	190.88	1%
Djuniord _{t-1}	21.66	1%
<hr/>		
Number of obs.	2266	
Number of banks	13	
Total gain	975.80	
Significance (total gain)	1%	

[†]Difference in normalized deviance between the GAM and the linear model. [‡]Based on a chi-square approximation to the distribution of the gain if the true marginal relationship between $sp_{i,b,t}$ and $x_{i,b,t}^j$ was linear.

are placed at the 33% and 67% percentile of each bank-specific distribution. This because the distributions of the regressors can vary greatly across banks, and placing knots without taking this into consideration may bias the analysis. For instance, all the observations for bank b could be in the upper interval of one variable because of structural factors and not because of a particular risk profile.

In order to reduce the number of coefficients to be estimated, I did not use splines for all the variables. In particular, the original specification is maintained for the control variables, for variables that are nonlinear by construction like $Dseniord_{t-1}$ and $Djuniord_{t-1}$ and for the assets variables (ta_{t-1} , $busrisk_t$, $\Delta busrisk_{t+4}$, and $liquidity_{t-1}$) since, according to the GAM model, nonlinearities play a less important role for this variable group.

In both tables, the coefficients reflect the slope in the respective interval. The results of the two models are quite different. The model with bank-specific knots appears to do a better job in explaining nonlinearities, especially for the performance variables. In Table 2.7, few coefficients are significant at at least the 5% level in more than one interval. Moreover, some variables, $provision_t$ and $\Delta provision_{t+4}$ in particular, have an unexpected sign. On the whole, however, the findings of both models support the view that bond prices are risk-sensitive, i.e., the market monitoring hypothesis cannot be rejected even after controlling for nonlinear effects.²⁷

To sum up, although the relationship between bond spreads and the vector of regressors is likely to be nonlinear, a nonlinear specification does not appear to provide additional insights compared to the basic model, also because it is very difficult to model the nonlinearities in a satisfactory manner. Although some convexities emerge, the basic conclusions concerning the main hypotheses remain valid. In other words, the linear model seems to be a relatively good first approximation of reality and the conclusions concerning the main hypotheses derived in the previous sections remain valid.

2.4 Conclusions

This study had two main objectives related to the feasibility of market discipline in the Swiss banking sector. First, I analyzed the scope for direct market discipline by testing the validity of one of its main prerequisites, namely that banks' investors are risk-sensitive and require adequate risk premia for the money they lend to banks. The results suggest that bondholders are rather skilled and forward-looking in assessing the soundness of banks. This is in line with the evidence provided in other countries. Yet, given the extremely low liquidity in the Swiss secondary debt market, this result was not obvious. The basic conclusions are also true for small banks, although these institutions do not publish detailed quarterly reports within the year. Debt issued by cantonal banks is an exception, however. Since these institutions enjoy a state guarantee, debt-holders have in principle no reason to monitor cantonal banks and to require any additional default risk premium.

²⁷Instead of using splines, one could also employ powers in order to control for nonlinear effects. I performed such a test, but the findings were very similar to those of the basic model.

Table 2.7: Spline model with general knots

Bank bond spreads are regressed on a set of explanatory variables using a PCSE model with bank-specific unobserved heterogeneity. For the variables modeled using linear splines, the knot points are given in parentheses. The sample consists of "normal" banks. All ratios are in percent.

Variable	Spline:		Low		Middle		High	
	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)
<i>Performance:</i>								
roa _t (0;0.5)	17.63	(14.55)	-49.36 [†]	(29.04)	-5.09	(8.43)		
Δroa _{t+4} (0;0.5)	11.90	(8.59)	-39.06*	(17.23)	9.72	(14.43)		
provision _t (3;6)	-10.44*	(5.89)	8.99*	(3.80)	-0.78	(3.14)		
Δprovision _{t+4} (-1;0)	3.90	(3.90)	-22.45**	(6.02)	5.81	(4.58)		
newprov _t (0.5;2)	-5.14	(20.37)	32.05**	(7.52)	-10.72 [†]	(5.57)		
npl _t (2;5)	1.51	(5.74)	7.40 [†]	(4.09)	-3.19	(2.67)		
Δnpl _{t+4} (-1;0)	4.92	(4.07)	39.46**	(7.66)	-15.00*	(5.98)		
<i>Assets:</i>								
ta _{t-1}	—	—	22.03*	(11.18)	—	—		
busrisk _t	—	—	0.20	(0.44)	—	—		
Δbusrisk _{t+4}	—	—	-0.11	(0.36)	—	—		
liquidity _{t-1}	—	—	0.02	(0.03)	—	—		
<i>SFBC rating:</i>								
ebk _t	—	—	16.65*	(7.88)	—	—		
Δebk _{t+1}	—	—	11.26	(10.14)	—	—		
<i>Liability structure:</i>								
deposit _{t-1} (20;40)	-1.39	(1.99)	-1.48	(1.74)	-3.43*	(1.29)		
seniord _{t-1} (10;20)	3.65*	(1.64)	-3.30*	(1.53)	6.90**	(2.54)		
juniord _{t-1} (1;2)	-17.90 [†]	(10.39)	30.52**	(11.47)	5.90	(9.98)		
Dseniord _{t-1}	—	—	0.98*	(0.27)	—	—		
Djuniord _{t-1}	—	—	-2.73	(2.39)	—	—		
Number of obs.	2266							
Number of banks	13							
R ²	0.77							
χ ²	7985.8							

Significance levels : †=10% *=5% **=1%

Table 2.8: Spline model with bank-specific knots

Bank bond spreads are regressed on a set of explanatory variables using a PCSE model with bank-specific unobserved heterogeneity. For the variables modeled using linear splines, the knot points are bank-specific. The sample consists of “normal” banks. All ratios are in percent.						
Variable	Spline:		Low		Middle	
	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)
<i>Performance:</i>						
roa _t	51.76**	(12.08)	8.60	(20.30)	33.83*	(14.56)
Δroa _{t+4}	-27.76**	(7.77)	34.33*	(17.02)	31.89**	(10.65)
provision _t	-4.22	(6.38)	-0.19	(4.13)	10.85*	(4.36)
Δprovision _{t+4}	11.21 [†]	(6.4)	-24.25**	(4.53)	13.67**	(5.11)
newprov _t	-36.19**	(10.58)	31.33**	(6.44)	19.00**	(4.3)
npl _t	-0.87	(4.55)	3.03	(5.76)	7.22**	(2.51)
Δnpl _{t+4}	12.31	(7.96)	22.45**	(4.72)	-1.10	(3.94)
<i>Assets:</i>						
ta _{t-1}	-	-	-16.76	(12.57)	-	-
busrisk _t	-	-	0.79*	(0.31)	-	-
Δbusrisk _{t+4}	-	-	0.43	(0.35)	-	-
liquidity _{t-1}	-	-	0.06*	(0.03)	-	-
<i>SFBC rating:</i>						
ebk _t	-	-	12.69	(8.48)	-	-
Δebk _{t+1}	-	-	16.70 [†]	(9.27)	-	-
<i>Liability structure:</i>						
deposit _{t-1}	-11.24**	(2.24)	2.94*	(1.28)	0.25	(1.62)
seniord _{t-1}	2.38	(1.74)	2.26	(2.43)	4.49*	(2.17)
juniord _{t-1}	-13.58	(12.51)	13.73	(12.86)	76.78**	(19.85)
Dseniord _{t-1}	-	-	0.99**	(0.27)	-	-
Djuniord _{t-1}	-	-	-3.68	(2.36)	-	-
<i>Summary statistics:</i>						
Number of obs.	2266					
Number of banks	13					
R ²	0.78					
χ ²	3228.8					
Significance levels :	†=10%	*=5%	**=1%			

Accordingly, I found that cantonal bond prices depend almost exclusively on systematic and market liquidity risk factors. In other words, this paper provides some additional empirical evidence that the state guarantee undermines market discipline. Some care is needed with this deduction, though, as no state-owned institution encountered major difficulties during the period under analysis.

The second aim of my study was to investigate whether the information embedded in bond yields may be useful to supervisors in identifying problems at the individual level, that is, if the market can indirectly discipline banks. I find little evidence that supervisors can learn something from the market. In particular, bondholders appear not to be able to discover potential sources of difficulties before the SFBC does. However, such a conclusion rests on the assumption that the SFBC does not make use of market data in its supervisory process.

In summary, there seems to be more scope for direct than for indirect market discipline in the Swiss banking sector. However, for direct market discipline to be effective, the risk-sensitive supply of funds must also affect the risk attitudes of banks. This issue has not yet been fully explored.

2.5 Appendix

2.5.1 Conditions for the acceptance of the debt contracts

Senior debt-holders will sign the contract if

$$E(S \mid \sigma^*) \geq r_f \lambda_S. \quad (2.13)$$

For a given σ the expected value of senior debt after signing the contract is

$$\begin{aligned} E(S \mid \sigma^*) &= E[\min(A, r^S \lambda_S) \mid \sigma^*] \\ &= \int_{\mu_{\sigma^*} - \sigma^*}^{r^S \lambda_S} A f(A \mid \sigma^*) dA + \int_{r^S \lambda_S}^{\mu_{\sigma^*} + \sigma^*} r^S \lambda_S f(A \mid \sigma^*) dA \\ &= \frac{1}{4\sigma^*} \left[2r^S \lambda_S (\mu_{\sigma^*} + \sigma^*) - (\mu_{\sigma^*} - \sigma^*)^2 - (r^S \lambda_S)^2 \right]. \end{aligned}$$

(2.13) is then satisfied if

$$2r^S \lambda_S (\mu_{\sigma^*} + \sigma^*) - (\mu_{\sigma^*} - \sigma^*)^2 - (r^S \lambda_S)^2 \geq 4\sigma^* r_f \lambda_S$$

or

$$r^S \geq \frac{\mu_{\sigma^*} + \sigma^* - 2[\sigma^* (\mu_{\sigma^*} - r_f \lambda_S)]^{\frac{1}{2}}}{\lambda_S}, \quad (2.14)$$

which holds as equality in equilibrium. By the same token, subordinated debt investors will sign the contract if

$$E(J \mid \sigma^*) \geq r_f \lambda_J \quad (2.15)$$

where

$$\begin{aligned} E(J \mid \sigma^*) &= E \left[\max(0, A - r^S \lambda_S) + \min(0, \tilde{\lambda} - A) \mid \sigma^* \right] \\ &= \int_{r^S \lambda_S}^{\tilde{\lambda}} (A - r^S \lambda_S) f(A \mid \sigma^*) dA + \int_{\tilde{\lambda}}^{\mu_{\sigma^*} + \sigma^*} r^J \lambda_J f(A \mid \sigma^*) dA \\ &= \frac{1}{4\sigma^*} \left[2r^J \lambda_J (\mu_{\sigma^*} + \sigma^*) + (r^S \lambda_S)^2 - \tilde{\lambda}^2 \right]. \end{aligned}$$

Hence, (2.15) holds if

$$2r^J \lambda_J (\mu_{\sigma^*} + \sigma^*) + (r^S \lambda_S)^2 - \tilde{\lambda}^2 \geq 4\sigma^* r_f \lambda_J$$

or

$$r^J \geq \frac{\mu_{\sigma^*} + \sigma^* - r^S \lambda_S - \left[(\mu_{\sigma^*} + \sigma^* - r^S \lambda_S)^2 - 4\sigma^* r_f \lambda_J \right]^{\frac{1}{2}}}{\lambda_J}.$$

Given the equilibrium value of r^S , this equation reduces to

$$r^J \geq \frac{2[\sigma^* (\mu_{\sigma^*} - r_f \lambda_S)]^{\frac{1}{2}} - 2[\sigma^* (\mu_{\sigma^*} - r_f \lambda_S - r_f \lambda_J)]^{\frac{1}{2}}}{\lambda_J}.$$

It is straightforward to show that both expressions in the square brackets are positive when $\mu_{\sigma^*} > r_f$, which holds by assumption.

Note that the total cost of debt in the case of no default amounts to

$$\begin{aligned} \tilde{\lambda} &= r^S \lambda_S + r^J \lambda_J \\ &= \mu_{\sigma^*} + \sigma^* - 2[\sigma^* (\mu_{\sigma^*} - r_f \lambda_S - r_f \lambda_J)]^{\frac{1}{2}} \end{aligned}$$

Hence, $\mu_{\sigma^*} > \tilde{\lambda}$ when

$$\begin{aligned} \sigma^* - 2[\sigma^* (\mu_{\sigma^*} - r_f \lambda_S - r_f \lambda_J)]^{\frac{1}{2}} &< 0 \\ \mu_{\sigma^*} - r_f \lambda_S - r_f \lambda_J &> \frac{\sigma^*}{4}. \end{aligned} \quad (2.16)$$

Since $\sigma_{\max} \leq \mu_{\sigma}$ by assumption, (2.16) holds as long as the expected return on assets is high enough when compared to the risk-free rate.

2 The prospects of market discipline for banks. Evidence from the Swiss bond market

A comparative static analysis shows that the equilibrium values of r^S and r^J react as expected to changes in their determinants. For instance, a higher asset risk implies a higher bond yield and hence a higher spread. Take first senior debt (μ constant):

$$\frac{\partial r^S}{\partial \sigma} = \frac{1}{\lambda_S} \left[1 - \left(\frac{\mu - r_f \lambda_S}{\sigma} \right)^{\frac{1}{2}} \right] > 0$$

by the assumption $\sigma_{\min} > \mu - r_f \lambda_S$. For subordinated bonds the same is true as

$$\frac{\partial r^J}{\partial \sigma} = \frac{1}{\lambda_J} \left[\left(\frac{\mu - r_f \lambda_S}{\sigma} \right)^{\frac{1}{2}} - \left(\frac{\mu - r_f \lambda_S - r_f \lambda_J}{\sigma} \right)^{\frac{1}{2}} \right] > 0.$$

Consider next the effect of μ for a given level of risk σ . From (2.14) one has

$$\frac{\partial r^S}{\partial \mu} = \frac{1}{\lambda_S} \left[1 - \left(\frac{\sigma}{\mu - r_f \lambda_S} \right)^{\frac{1}{2}} \right] < 0.$$

By the same token

$$\frac{\partial r^J}{\partial \mu} = \frac{1}{\lambda_J} \left[\left(\frac{\sigma}{\mu - r_f \lambda_S} \right)^{\frac{1}{2}} - \left(\frac{\sigma}{\mu - r_f \lambda_S - r_f \lambda_J} \right)^{\frac{1}{2}} \right] < 0.$$

Let me finally examine the effect of leverage (σ constant). Consider first r^S :

$$\frac{\partial r^S}{\partial \lambda_S} = \frac{r_f \lambda_S \left(\frac{\sigma}{\mu - r_f \lambda_S} \right)^{\frac{1}{2}} - \mu - \sigma + 2[\sigma(\mu - r_f \lambda_S)]^{\frac{1}{2}}}{(\lambda_S)^2}.$$

$\frac{\partial r^S}{\partial \lambda_S} > 0$ when

$$\begin{aligned} \sigma r_f \lambda_S - (\mu + \sigma) [\sigma(\mu - r_f \lambda_S)]^{\frac{1}{2}} + 2\sigma(\mu - r_f \lambda_S) &> 0 \\ \sigma \left(\frac{r_f \lambda_S}{\mu - \sigma} \right)^2 &> \mu - r_f \lambda_S \end{aligned}$$

which holds by $\sigma_{\min} > \mu - r_f \lambda_S$. Take now r^J . More senior debt implies a higher subordinated debt spread since

$$\frac{\partial r^J}{\partial \lambda_S} = \frac{1}{\lambda_J} \left[r_f \left(\frac{\sigma}{\mu - r_f \lambda_S - r_f \lambda_J} \right)^{\frac{1}{2}} - r_f \left(\frac{\sigma}{\mu - r_f \lambda_S} \right)^{\frac{1}{2}} \right] > 0.$$

Define $\eta = [4\sigma(\mu - r_f \lambda_S)]^{\frac{1}{2}}$. Then

$$\frac{\partial r^J}{\partial \lambda_J} = \frac{1}{(\lambda_J)^2} \left[2\sigma r_f \lambda_J (\eta^2 - 4\sigma r_f \lambda_J)^{-\frac{1}{2}} - \eta + (\eta^2 - 4\sigma r_f \lambda_J)^{\frac{1}{2}} \right]$$

which is greater than zero since

$$\begin{aligned}
2\sigma r_f \lambda_J - \eta (\eta^2 - 4\sigma r_f \lambda_J)^{\frac{1}{2}} + \eta^2 - 4\sigma r_f \lambda_J &> 0 \\
\eta^2 - 2\sigma r_f \lambda_J &> \eta (\eta^2 - 4\sigma r_f \lambda_J)^{\frac{1}{2}} \\
(\eta^2 - 2\sigma r_f \lambda_J)^2 &> \eta^4 - 4\eta^2 \sigma r_f \lambda_J \\
(2\sigma r_f \lambda_J)^2 &> 0.
\end{aligned}$$

2.5.2 Expected value of equity in the infinitely repeated game

In an infinitely repeated game the managers maximize

$$E_{t-1}(C_t) = E_{t-1} \left(\sum_{i=0}^{\infty} \frac{D_{t+i}}{r_f^i} \right).$$

As in the one-period model, at the beginning of the game the expected value of the assets at the end of period $t = 1$ – given that the game did not end at time $t = 1$ – is given by

$$E_0(A_1 \mid A_1 > \tilde{\lambda}) = \frac{\mu_\sigma + \sigma + \tilde{\lambda}}{2}.$$

Since

$$\begin{aligned}
&E_0[A_2 \mid A_1 > \tilde{\lambda}, A_2 > \tilde{\lambda}] \\
&= E_0[E_1(A_2 \mid A_2 > \tilde{\lambda}) \mid A_1 > \tilde{\lambda}] \\
&= E_0\left[\frac{\mu_\sigma + \sigma + \tilde{\lambda}}{2} \mid A_1 > \tilde{\lambda}\right] \\
&= \frac{\mu_\sigma + \sigma + \tilde{\lambda}}{2}
\end{aligned}$$

and therefore

$$E_0(A_t \mid \text{Game did not end at } t) = \frac{\mu_\sigma + \sigma + \tilde{\lambda}}{2}.$$

Moreover, the probability at time $t = 0$ that the game is still “alive” at time $t + 1$ is given by

$$\begin{aligned}
Pr_0(\text{Game did not end at } t) &= \prod_{i=1}^t Pr_0(A_i > \tilde{\lambda} \mid \text{Game did not end at time } i-1) \\
&= Pr_0(A_1 > \tilde{\lambda}) \cdot Pr_0(A_2 > \tilde{\lambda} \mid A_1 > \tilde{\lambda}) \cdot \dots \\
&= \left(\frac{\mu_\sigma + \sigma - \tilde{\lambda}}{2\sigma} \right)^t \\
&= \pi(\sigma)^t.
\end{aligned}$$

2 The prospects of market discipline for banks. Evidence from the Swiss bond market

For $\sigma \in [\sigma_{\min}; \sigma_{\max}]$ and $\mu_\sigma > \tilde{\lambda}$ the restrictions $0 \leq \pi(\sigma) \leq 1$ hold. In fact, given that $\pi(\sigma)$ is monotonically decreasing in σ ,

$$\begin{aligned}\pi(\sigma_{\min}) &= \frac{\mu_\sigma + \mu_\sigma - r_f \lambda_S - \tilde{\lambda}}{2(\mu_\sigma - r_f \lambda_S)} = 1 - \frac{\tilde{\lambda} - r_f \lambda_S}{2(\mu_\sigma - r_f \lambda_S)} \leq 1 \\ \pi(\sigma_{\max}) &= \frac{\mu_\sigma + \mu_\sigma - \tilde{\lambda}}{2\mu_\sigma} = 1 - \frac{\tilde{\lambda}}{2\mu_\sigma} \geq 0.\end{aligned}$$

The expected value of the first dividend equals

$$\begin{aligned}E_0(D_1) &= E_0 \left[A_1 - 1 - (r^S - 1) \lambda_S - (r^J - 1) \lambda_J \mid A_1 > \tilde{\lambda} \right] \\ &\quad \cdot Pr_0(A_1 > \tilde{\lambda}) \\ &= \left[\frac{\mu_\sigma + \sigma + \tilde{\lambda}}{2} - 1 - (r^S - 1) \lambda_S - (r^J - 1) \lambda_J \right] \\ &\quad \cdot \frac{\mu_\sigma + \sigma - \tilde{\lambda}}{2\sigma} \\ &= \left(\frac{\mu_\sigma + \sigma - \tilde{\lambda}}{2} - 1 + \lambda_S + \lambda_J \right) \frac{\mu_\sigma + \sigma - \tilde{\lambda}}{2\sigma} \\ &= \delta(\sigma) \pi(\sigma)\end{aligned}$$

and

$$\begin{aligned}E_0(D_t) &= E_0 \left[A_t - 1 - (r^S - 1) \lambda_S - (r^J - 1) \lambda_J \mid A_t > \tilde{\lambda} \right] \\ &\quad \cdot Pr_0(\text{Game did not end at } t) \\ &= \delta(\sigma) \pi(\sigma)^t.\end{aligned}$$

Bringing all together

$$\begin{aligned}E_{t-1}(C_t) &= E_{t-1} \left(\sum_{i=0}^{\infty} \frac{D_{t+i}}{r_f^i} \right) \\ &= \sum_{i=0}^{\infty} \frac{\delta(\sigma) \pi(\sigma)^{i+1}}{r_f^i} \\ &= \delta(\sigma) \pi(\sigma) \sum_{i=0}^{\infty} \left[\frac{\pi(\sigma)}{r_f} \right]^i.\end{aligned}$$

Since $\pi(\sigma) \leq 1$ and $r_f > 1$ by assumption, one finally gets

$$E_{t-1}(C_t) = \frac{\delta(\sigma) \pi(\sigma) r_f}{r_f - \pi(\sigma)}.$$

Suppose now μ_σ is constant. With a little algebra one can show that $E_{t-1}(C_t)$ is a convex function of σ . A sufficient condition for convexity is that

$$\frac{\partial^2 E_{t-1}(C_t)}{\partial \sigma^2} = \frac{4r_f^2 (\mu - \tilde{\lambda}) M}{(\tilde{\lambda} - \mu - \sigma + 2r_f \sigma)^3} > 0 \quad (2.17)$$

with

$$M = 1 - \lambda_S - \lambda_J - r_f (2 + \tilde{\lambda} - 2\lambda_S - 2\lambda_J - \mu).$$

The denominator of (2.17) is positive. To see this, note first that it is monotonically increasing in σ and $\tilde{\lambda}$. So it suffices to prove that $r^S \lambda_S + r^J \lambda_J - \mu - \underline{\sigma} + 2r_f \underline{\sigma} > 0$, with $\underline{\sigma} = \mu - r_f \lambda_S < \sigma_{\min}$ and $r^S = r^J = r_f < r^S$. By substituting one gets

$$\begin{aligned} & r_f (\lambda_S + \lambda_J) - 2\mu + r_f \lambda_S + 2r_f (\mu - r_f \lambda_S) \\ &= 2\mu (r_f - 1) + r_f \lambda_J - 2r_f \lambda_S (r_f - 1) \\ &= 2(r_f - 1) (\mu - r_f \lambda_S) + r_f \lambda_J > 0. \end{aligned}$$

So if M is positive, $\frac{\partial^2 E_{t-1}(C_t)}{\partial \sigma^2}$ is also positive. Unfortunately, one cannot easily demonstrate that this is always true. However, consider the extreme case with $r^S = r^J = \mu$. Then M reduces to

$$(1 - \lambda_S - \lambda_J) (1 - 2r_f + r_f \mu)$$

which is always greater than 0 for $\mu > 2 - \frac{1}{r_f}$. When $r_f > 1$ this is always satisfied given the assumption $\mu > r_f$. Hence, since M decreases in $\tilde{\lambda}$ and both bond yields are lower than μ under normal circumstances, $E_{t-1}(C_t)$ is a convex function of σ .

2.5.3 Introducing concavity into the manager's objective function

Suppose $\mu_\sigma = \mu_M - \kappa (\sigma - \sigma_M)^2$. Moreover $\mu_M = 1.2$, $\sigma_M = 0.9$, $\lambda_S = 0.5$, $\lambda_J = 0$ and $r_f = 1.05$. Figure 2.2 shows the shape of the objective function of the manager in the one-period game (2.3) for several values of r^S and κ . It is easy to show that for $\kappa = 1$ and $\kappa = 2$, $\{\sigma = 0.989; r^S = 1.113\}$ and $\{\sigma = 0.938; r^S = 1.094\}$ respectively represent the subgame perfect equilibrium of this game. In both cases the expected return on equity is greater than r_f .

Figure 2.3 shows the shape of the objective function of the manager in the infinitely repeated game (2.8). Again, $\{\sigma = 0.844; r^S = 1.07\}$ represents the subgame perfect equilibrium with $\kappa = 2$. Note that in the dynamic game, the expected payoff for the equity-holder is on average higher than in the static case, while the optimal level of risk is lower. This because the bank has a higher franchise value that reduces the incentives for taking risks.

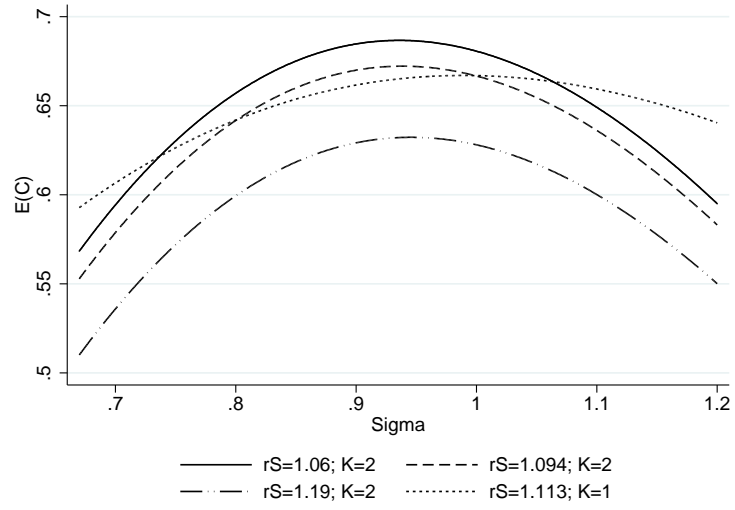


Figure 2.2: Objective function of the manager in the static model

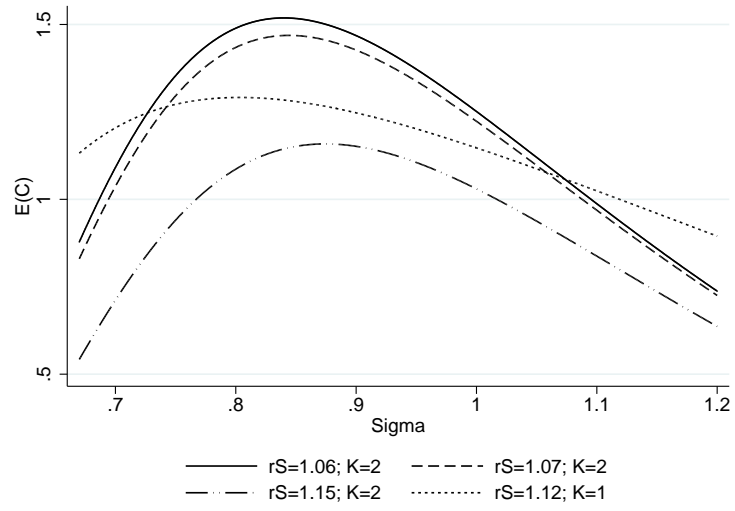


Figure 2.3: Objective function of the manager in the dynamic model

2.5.4 Variable definition

Table 2.9 describes the variables used in the estimations as well as their availability and frequency of data collection respectively.

When possible, for UBS and CSG figures on a consolidated basis are used. For all other institutions I employ data at bank level, since subsidiaries play a minor role as opposed to the two big banks.

All bank book data come from the internal database of the Swiss National Bank.

Table 2.9: Variable description

Name	Description	Frequency	Available since
$sp_{b,t}$	Bond spread (in basis points)	Monthly	January 1990
$roa_{b,t}$	Net income to total assets	Yearly	1990
$provisions_{b,t}$	Current provisions to total credits	Yearly	1995
$newprov_{b,t}$	Net new provisions to total credits	Yearly	1998
$npl_{b,t}$	Non-performing loans to total credits	Yearly	1998
$ta_{b,t}$	Log of total assets in million Swiss francs	Monthly	January 1990
$busrisk_{b,t}$	Regulatory risk-weighted assets to total assets	Quarterly	December 1997
$liquidity_{b,t}$	Bank liquidity (quick ratio)	Quarterly	March 1990
$ebk_{b,t}$	“SFBC-Rating” (1 if problem bank; 0 otherwise)	Monthly	January 1990
$deposit_{b,t}$	Amount of deposits to total assets	Monthly	January 1990
$senior_{b,t}$	Senior debt outstanding to total assets	Monthly	January 1990
$junior_{b,t}$	Subordinated debt outstanding to total assets	Monthly	January 1990
$Dsenior_{b,t}$	$senior_{b,t}$ if subordinated bond; 0 otherwise	Monthly	January 1990
$Djunior_{b,t}$	$junior_{b,t}$ if senior bond; 0 otherwise	Monthly	January 1990
$gdpg_{b,t}$	Annual GDP growth rate (seasonally adjusted)	Quarterly	March 1990
$unempl_t$	Unemployment rate (seasonally adjusted)	Monthly	January 1990
ird_t	Interest rate differential for whole bank sector	Yearly	1990
$riskfree_t$	1 year government bond rate	Monthly	January 1990
$term_t$	Term structure factor	Monthly	January 1990
$default_t$	Default premium factor	Monthly	January 1990
$market_t$	Market excess return over the risk-free rate	Monthly	January 1990
smb_t	Small minus big size factor	Monthly	May 1993
hml_t	High minus low book-to-market factor	Monthly	January 1990
$volimpl_t$	Leu implied volatility index	Monthly	May 1995
$amount_i$	Bond's total amount outstanding	Monthly	–
$coupon_i$	Bond's coupon	Monthly	–
$mtm_{i,t}$	Months until bond's maturity	Monthly	–
sub_i	1 if subordinated bond; 0 otherwise	Monthly	January 1990

2.5.5 The variance covariance matrix

$$\Omega = \begin{bmatrix} \Omega_{11} & \Psi_1 & \dots & 0 & \dots & \dots & 0 \\ \Psi_1 & \Omega_{21} & & & & & \dots \\ \dots & & \dots & & & & \dots \\ 0 & & & \Omega_{bi} & & & 0 \\ \dots & & & & \dots & & \dots \\ & & & & & \Omega_{I-1B} & \Psi_B \\ 0 & \dots & \dots & 0 & \dots & \Psi_B & \Omega_{IB} \end{bmatrix}$$

with

$$\Omega_{ib} = \begin{bmatrix} \tilde{\sigma}_b^2 + \hat{\sigma}_{i,b}^2 + \sigma^2 & \tilde{\sigma}_b^2 + \hat{\sigma}_{i,b}^2 & \dots & \tilde{\sigma}_b^2 + \hat{\sigma}_{i,b}^2 \\ \tilde{\sigma}_b^2 + \hat{\sigma}_{i,b}^2 & \dots & & \dots \\ \dots & & \dots & \tilde{\sigma}_b^2 + \hat{\sigma}_{i,b}^2 \\ \tilde{\sigma}_b^2 + \hat{\sigma}_{i,b}^2 & \dots & \tilde{\sigma}_b^2 + \hat{\sigma}_{i,b}^2 & \tilde{\sigma}_b^2 + \hat{\sigma}_{i,b}^2 + \sigma^2 \end{bmatrix}$$

for $b = 1, \dots, B$ and $i \in I_b$ and

$$[\Psi_b]_{mn} = \tilde{\sigma}_b^2$$

for every m, n . I_b is the set of bonds issued by bank b .

2.5.6 Complete estimation results

The complete estimation results for sample A (“normal” banks) and sample B (insured cantonal banks) are shown in Table 2.10.

Table 2.10: Complete base-model estimation results

Bank bond spreads are regressed on a set of explanatory variables using a PCSE model with bank-specific unobserved heterogeneity. All ratios are in percent.				
Variable	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)
Sample:	(A) "Normal" banks		(B) Insured banks	
roa _t	15.79**	(3.90)	19.17	(13.31)
Δ roa _{t+4}	5.08 [†]	(2.76)	-1.34	(14.05)
provision _t	3.13 [†]	(1.72)	10.56**	(2.71)
Δ provision _{t+4}	2.63	(2.35)	4.81 [†]	(2.54)
newprov _t	13.21**	(2.33)	-5.52	(6.31)
npl _t	-1.19	(1.69)	-0.86	(1.27)
Δ npl _{t+4}	11.92**	(2.62)	-3.52**	(1.10)
ta _{t-1}	-3.48	(8.30)	29.46	(34.98)
busrisk _t	1.05**	(0.29)	-0.15	(0.37)
Δ busrisk _{t+4}	0.34	(0.33)	0.59 [†]	(0.32)
liquidity _{t-1}	0.04	(0.03)	-0.01	(0.01)
ebk _t	18.50**	(7.09)	1.04	(9.74)
Δ ebk _{t+1}	15.68	(9.64)	27.98 [†]	(14.55)
deposit _{t-1}	0.19	(0.74)	1.33	(0.86)
seniord _{t-1}	2.03*	(0.94)	-0.62 [†]	(0.36)
juniord _{t-1}	10.42	(6.74)	-11.15	(7.51)
Dseniord _{t-1}	1.01**	(0.26)	-0.20	(0.43)
Djuniord _{t-1}	-2.34	(2.32)	8.12	(7.82)
ird _{t+4}	-149.24*	(64.52)	-44.42	(57.06)
gdpgr _{t+1}	2.10	(3.67)	2.29	(2.73)
unempl _{t+1}	45.23**	(12.26)	35.34**	(9.85)
riskfree _{t-1}	24.50**	(8.72)	27.47**	(7.13)
term _{t-1}	7.02	(5.73)	-2.5	(4.59)
default _{t-1}	63.15**	(8.87)	32.06	(7.10)
volimpl _{t-1}	1.17**	(0.26)	0.58**	(0.22)
market _{t-1}	0.45 [†]	(0.24)	0.36*	(0.19)
smb _{t-1}	-0.18	(0.16)	-0.34*	(0.13)
hml _{t-1}	0.08	(0.41)	0.13	(0.31)
amount	0.01	(0.01)	-0.01	(0.01)
coupon	2.67*	(1.14)	1.44	(0.88)
mtm _t	0.30**	(0.05)	0.25**	(0.04)
mtmr _t	-10.04 [†]	(5.96)	-7.42 [†]	(4.30)
sub	24.38**	(4.97)	42.44*	(16.84)
constant	-58.99	(128.77)	-430.09	(324.59)
Number of obs.	2266		2841	
Number of banks	13		17	
R ²	0.74		0.45	
χ^2	17672.0		3089.2	
Significance levels :	†=10%	*=5%	**=1%	

3 Is direct market discipline in banking really effective and necessary? A comparison between public and private banks

A substantial amount of research has yielded evidence that market data – prices or quantities – contain relatively accurate information about the risk profile of corporations, and banks in particular.¹ Bond spreads, for instance, tend to increase when the creditworthiness of the issuer deteriorates. This means that investors such as depositors or bondholders are risk-sensitive and require an adequate risk premium for the money they lend to banks. Following the distinction between market monitoring and market influence emphasized by Bliss and Flannery (2002), this paper analyzes whether such a risk-sensitive supply of funds has an effect on bank behavior, i.e., whether the market can directly discipline banks in that it reduces managers' incentives for behaving “improperly” and assuming bad or excessive risks.

I test whether direct market discipline is really effective by comparing the lending policy and performance of public banks – which should not be subject to market discipline as they enjoy a state guarantee – with those of private banks.² As illustrated by Merton (1977) in the context of deposit insurance, safety nets have moral hazard consequences in that they increase the risk appetite of equity-holders above efficient levels, even in the absence of bankruptcy costs. This is especially true when safety nets are free or priced independently of risk. Demirgüç-Kunt and Detragiache (2002) and Demirgüç-Kunt and Huizinga (2004) provide cross-country empirical evidence for the distorting effects of deposit insurance. Given that state guarantees go beyond deposit insurance and often cover the whole debt of a bank, moral hazard could be even worse at public banks.³ As a result, public institutions may implement riskier strategies than private banks that are subject to direct market discipline, and allocate their resources inefficiently, which generates costs for the taxpayer. On the other hand, without market discipline public and private banks are expected to

¹See Flannery (1998) for an overview and DeYoung, Flannery, Lang, and Sorescu (2001), Swidler and Wilcox (2002), Fan, Haubrich, Ritchken, and Thomson (2003), Krainer and Lopez (2004), Krishnan, Ritchken, and Thomson (2006) or Curry, Elmer, and Fissel (2007) for more recent research on the US case. Persson (2002), Sironi (2003), Cannata and Quagliariello (2005) and Gropp, Vesala, and Vulpes (2006) provide evidence for European banks, Birchler and Maechler (2002) and Facchinetti (2007) for Swiss banks.

²Throughout the paper the term “private bank” refers to all those institutions that do not benefit from any public support. It should be noted that in Switzerland the notion of “private bank” is usually associated with an institution specialized in off-balance-sheet asset management.

³In some cases, however, states require a risk-adjusted compensation for the guarantee (see Section 3.4.2). This should a priori reduce moral hazard.

3 Is direct market discipline in banking really effective and necessary?

behave more similarly, as moral hazard affects the decision process of both bank types. The Swiss banking sector offers an ideal background for testing this hypothesis empirically. In fact, cantonal banks constitute Switzerland's second biggest group of banks. To a large degree, they are public institutions whose liabilities are guaranteed by the respective canton.⁴ Most of them are universal banks with a strong focus on savings and mortgage business in their respective region. My idea is to compare the credit policy and performance of cantonal banks with that of other Swiss banks with a similar business model.

There is a fundamental issue with respect to this approach, though. Moral hazard arises when managers act on behalf of equity-holders and maximize shareholder value and in particular the benefits that arise from limited liability. While this is a standard assumption for private banks, a public bank and its major equity-holder, the state, may have other objectives. In the absence of agency costs, it is actually inconsistent to believe that managers exploit the limited liability status of equity, given that the equity-holder itself assures all liabilities! This issue is closely connected to the more general question on the reasons for the existence of public banks in developed countries.⁵ One could argue that the purpose of the state guarantee is to avoid systemic threats such as self-fulfilling depositor runs. However, this goal may be also achieved by a simple deposit insurance scheme, as in Diamond and Dybvig (1983). Therefore, a full guarantee must have some other goal, otherwise its "added value" would be just an indirect subsidy of the private equity-holders (if any) of public banks. The typical public mandate stipulates that the cantonal bank should support the local economy and local business enterprises. What this actually means in reality is a priori not clear. As a result, if one wants to compare the credit policies of private and public banks from the perspective of market discipline, assumptions regarding the objective function of public banks are needed.

In addition, one has to control for the effects of the franchise or charter value, defined as the present value of the bank as an ongoing institution. Charter value may dampen or even eliminate the moral hazard problem caused by safety nets or by the absence of market discipline (Buser, Chen, and Kane, 1981; Marcus, 1984; Keeley, 1990; Marshall and Prescott, 2001). Intuitively, when the shareholders or the managers of a bank have something to lose in the case of default, in particular expected future income, risky business becomes less attractive. A number of papers (see, for instance, Keeley, 1990; Demsetz, Saidenberg, and Strahan, 1996; Gropp and Vesala, 2001; González, 2005) show that this hypothesis is consistent with cross- and within-country empirical evidence. In the extreme, i.e., when the charter value is so high that banks just want to minimize risk, the concept of market discipline, and banking regulation more generally, would make less

⁴Switzerland consists of 26 cantons (federal states).

⁵La Porta, Lopez-De-Silanes, and Shleifer (2002) discuss two basic views that can explain the existence of state-owned banks. First, when private banks are not sufficiently developed, only public banks may be able to provide the necessary funds needed to foster economic growth. Yet, this optimistic view can be hardly applied to the modern, highly developed Swiss banking sector. The second rationale is that politicians may see public banks as instruments for pursuing personal objectives – the "political view". For additional evidence on the role of public banks in non-industrialized countries, see Andrews (2005) or De Nicolò and Loukoianova (2006).

3.1 Related literature

sense and only if banks' defaults have substantial negative externalities would there be any reason to further limit risk-taking. On the other hand, the state guarantee may prevent banks from being excessively risk-averse and thus possibly generate economic value added, which would justify the existence of public banks in developed countries.

To account for those factors, I first develop a simple theoretical model that allows me to analyze the impact of market discipline and public bank's objectives on the lending policy of private and public banks. I show that in a static world with direct market discipline private banks invest efficiently, while moral hazard or less effective monitoring (or both) may induce the managers of public banks to assume bad risks. On the other hand, when public banks pursue other objectives such as the maximization of economic value added, the lending strategies of the two bank groups become more similar. In the absence of market discipline, a private bank fully exploits the limited liability status of equity and hence it tends to be more aggressive than a public bank, independently of the objective function of the latter. In a more dynamic context, the value of the franchise becomes a critical determinant of the behavior of private bank managers: the higher the charter value, the lower the incentives for the private bank to assume risks. For a public bank, the franchise value generally plays a minor role. If the bank is always bailed out, it can be assumed that there is no danger of losing the franchise value.

In spite its simplicity, the model allows me to derive sensible empirical implications for banks' lending performance, which I test using data on Swiss cantonal banks, regional banks and the Raiffeisen group. All banks in the sample are mainly active in the regional Swiss mortgage and loan business and are subject to the same regulatory and legal system. Thus, they represent a useful basis for comparison for the purposes of this paper. Possible cross-sectional differences in lending performance can be hardly explained by environmental factors, which are not always easy to control for. This is a major advantage of this paper when compared to other studies, which often rely on cross-country panels.

I find strong evidence that cantonal banks assume more credit risks than other bank groups. Hence, they tend to generate larger gross interest revenues. However, those extra returns do not cover the additional losses caused by the higher default risk. This result can be hardly justified by arguing that cantonal banks are more customer-friendly when setting out the terms of the loan contracts. In other words, moral hazard and/or less strict monitoring by the state as equity-holder appear to be a problem for cantonal banks and can lead to an inefficient allocation of resources. More importantly, my results suggest that the market – e.g., depositors – exercises some discipline on private banks and even on those that do not issue securities on the capital market (the typical channel through which discipline is expected to work). My findings also suggest that the franchise value hardly affects the portfolio choice of the typical Swiss regional bank.

The paper is organized as follows: Section 3.1 relates the paper to the existing literature. Section 3.2 introduces the static model and derives its empirical implications. Section 3.3 discusses the basic issues that arise in a more dynamic framework. In Section 3.4, I present the empirical setup. Estimation results are shown in Section 3.5. Section 3.6 concludes.

3.1 Related literature

Several researchers have compared the performance of public banks with the performance of other bank groups in industrial countries. Rime and Stiroh (2003) examine scale economies, scope economies and profit-efficiency of Swiss banks between 1996 and 1999, the main aim being to explain the consolidation process in the Swiss banking industry. Among other things, their results reject the hypothesis that cantonal banks are more cost- and profit-efficient than other Swiss bank groups and regional banks in particular. In similar studies focusing on the German and Austrian banking sectors, Altunbas, Evans, and Molyneux (2001) and Hauner (2005) find that German and Austrian public banks are more efficient, but they attribute the result to the lower cost of funds enjoyed by public institutions. Hauner argues that this discount may be attributable to the state guarantee. Bichsel (2006) analyzes the role of cantonal banks from the perspective of competition and concludes that they do not deviate from the typical conduct in the Swiss banking sector. In addition, he provides empirical evidence showing that cantonal banks have lower-than-average funding costs by looking at interest rates on deposits. Bichsel suggests that this discount may be explained by the state guarantee. Birchler and Maechler (2002) examine whether depositors withdraw their savings deposits when the risk profile of their bank deteriorate and find that depositors of cantonal banks are less responsive to changes in bank fundamentals than those of private banks. Sironi (2002, 2003), Pop (2006) and Facchinetti (2007) arrive at similar conclusions in the context of capital market-based funding. Facchinetti provides evidence that spreads of bonds of cantonal banks contain little idiosyncratic information on the risk profile of the issuer. By looking at subordinated debt issues in Europe, Sironi (2003) observes that, after controlling for the bank's risk, bonds issued by public banks trade at an average discount premium of about 40 basis points. Moreover, Sironi's findings also support the hypothesis that debt investors are able to discriminate between the different risk profiles of private banks, but they care less about risk when pricing bonds issued by public or guaranteed institutions. Sironi (2002) and Pop argue that because of the state guarantee, implicit or explicit, the average European and Japanese bank bond trades at a lower spread than its North American counterpart. Finally, Flannery and Sorescu (1996) show that US bank investors became risk-sensitive only when the regulator made clear that they were not protected in the case of default. Consequently, one can conclude that in the presence of state guarantees investors hardly price idiosyncratic risk and hence a prerequisite of direct market discipline is not satisfied. However, there is almost no study that goes one step further and examines whether this risk-insensitive supply of funds induces public banks to be less cautious in lending money than other bank groups. One major exception is the paper by Iannotta, Nocera, and Sironi (2007). The authors compare risk and performance of European public, mutual and private banks and find that public banks are, on average, less profitable but at the same time they tend to have lower quality assets. In Switzerland this issue is of particular relevance considering the importance of cantonal banks for the Swiss banking sector. Furthermore, given the evidence that cantonal banks have lower financing costs thanks to the state guarantee, the fact that they are not more

profit-efficient than other Swiss banks (Rime and Stiroh, 2003) is left unexplained.

Another branch of the literature examines the empirical relevance of direct market discipline, especially in the US. Those studies typically focus on the relationship between bank risk and (subordinated) debt contracts. The evidence from this literature is mixed. Bliss and Flannery (2002) analyze whether bank holding companies' stock and bond prices affect a set of management actions⁶ but their results provide little support for investor influence. Goyal (2005), on the other hand, observes that banks with greater risk-taking incentives – which are determined by the bank's franchise value – include more restrictive covenants in their debt contracts, which should reduce yields spreads but also managerial flexibility to increase risk. Nier and Baumann (2006) find a positive relationship between bank equity capital and several proxies for the strength of market discipline (depositor protection, uninsured funding and transparency) in a panel of banks from different countries. Yet, much less work has been done on analyzing the effects of market discipline directly on the asset side of the balance sheet.

In this contribution, I try to close all these gaps and link these two branches of the literature.

3.2 A simple static model

In this section I model bank lending behavior in a very simple way. Assume an economy that comprises one bank as well as a number of equity and debt investors. Everybody is risk neutral. Debt holders have two investment opportunities. On the one hand, they can lend money to the bank. In this case the liability side of the bank's balance sheet will consist of debt D and equity E . The bank then has the option of investing its assets $A = D + E$ in a credit portfolio. On the other hand, if debt investors decide not to finance the bank, the latter will be liquidated and everybody receives a risk-free gross return r_f .

The bank can choose between two credit portfolio types: a risky portfolio, which I call the H-portfolio (High risk), and a relatively safer one, the L-portfolio (Low risk). As an example, the H-portfolio may consist of unsecured loans to small and medium enterprises (SMEs), while residential mortgages could make up the L-portfolio. Table 3.1 gives a more accurate description of the payoffs of each portfolio:

Table 3.1: Return distribution of the portfolios

	H-portfolio	L-portfolio
Probability of no default	π_H	π_L
Cash flow if no default	R_H	R_L
Cash flow if default	0	r_L

I assume that the probability of default of the riskier portfolio is higher, i.e., $\pi_L > \pi_H$, but

⁶This set includes several factors affecting bank leverage (e.g., increase in equity as percentage of book value of equity), one factor affecting asset risk (the change in securities portfolio as a proportion of total assets) and other measures.

3 Is direct market discipline in banking really effective and necessary?

$\pi_H R_H > \pi_L R_L$ otherwise the L-portfolio would a priori always dominate the H-portfolio in net present value (NPV) terms. Moreover, $\pi_L R_L + (1 - \pi_L) r_L > r_f A$ and $\pi_H R_H > r_f A$, so that both credit portfolios have a positive NPV. Let $r_L < D$ so that if either portfolio defaults, the bank becomes insolvent. Debt enjoys a seniority status in the case of default. Note that if both portfolios have the same NPV and bankruptcy costs are positive, the L-portfolio is preferred by a hypothetical regulator for two reasons. First, the probability of default is smaller. Second, in the case of default the loss is higher for the risky portfolio.

Gross risk-free interest rate and A are normalized to one for simplicity.

3.2.1 Market discipline

It is assumed that market discipline – defined here as a risk-sensitive supply of debt – may only affect the lending policy of a private bank, since lenders of public banks are protected by the guarantee and hence do not require any risk premium. In other words, public banks can borrow at the risk-free rate. In what follows, I show that in the presence of market discipline a private bank invests efficiently. Public banks, on the other hand, may be too risk-aggressive. This is especially true when the state does not exercise sufficient control over the management.

Private bank

Consider first a private bank that does not benefit from any state guarantee or liability insurance scheme. If the bank is insolvent, it will be liquidated (no bail-out) at no additional costs.⁷ Suppose, moreover, that debt investors are well-informed about the investment opportunities of the bank and hence know in which portfolio $i \in \{H, L\}$ the managers invest. Accordingly, they require an adequate compensation c_i for lending money, which depends on the portfolio chosen by the managers (market discipline). Therefore D is not risk-free. The management maximizes shareholder value; that is, I assume zero agency costs between managers and equity-holders. In a model about market discipline and moral hazard, this may seem odd at first sight. However, the general issue of conflicts of interest at private firms is beyond the scope of this paper and hence – at least with respect to private banks – assuming no agency costs can be considered as a standard assumption.⁸ Note, moreover, that market discipline as defined in this paper should be viewed as an instrument aimed at reducing moral hazard that arises from the limited liability status of equity and not from agency costs. After all, if even equity-holders were not able to accurately monitor managers' activities, one should strongly question the mere existence of external market monitoring and discipline.

In the absence of agency costs the management will choose the project that generates the highest expected return after borrowing costs, provided that this return exceeds the risk-free rate. The

⁷Another possibility is that the bank may be rescued and restructured, such that debt holders are not compensated for their losses.

⁸See Altunbas, Evans, and Molyneux (2001) for a discussion of the “standard” conflict of interest between owner and managers in the banking industry.

H-portfolio represents an optimal investment if

$$\begin{aligned} \pi_H (R_H - c_H D) &\geq \pi_L (R_L - c_L D) \\ \text{s.t. } \pi_H (R_H - c_H D) &\geq E, \end{aligned}$$

i.e., if

$$\pi_H R_H - \pi_L R_L - (\pi_H c_H - \pi_L c_L) D \equiv \Psi \geq 0 \quad (3.1)$$

$$\text{s.t. } \pi_H R_H - 1 + (1 - \pi_H c_H) D \geq 0. \quad (3.2)$$

(3.1) basically states that for the H-portfolio to be optimal, the expected additional revenues must exceed the upsurge in the borrowing costs. Note that r_L does not play any direct role in the decision of the managers, since in the case of default the equity-holders receive nothing in any case. r_L affects the borrowing costs, though. But what are the values of c_H and c_L for which debt holders are willing to finance the bank? The investors will lend money to the bank only if the expected return on the debt, which depends on the portfolio chosen by the managers, is higher than or equal to r_f . Hence, c_H solves

$$\begin{aligned} \pi_H c_H D &\geq D \\ c_H &\geq \frac{1}{\pi_H} \end{aligned} \quad (3.3)$$

whereas c_L is determined by

$$\begin{aligned} \pi_L c_L D + (1 - \pi_L) r_L &\geq D \\ c_L &\geq \frac{1}{\pi_L} \left[1 - (1 - \pi_L) \frac{r_L}{D} \right] \end{aligned} \quad (3.4)$$

I assume that, when indifferent, the investors lend the money to the bank and that the managers have all the bargaining power. As a result (3.3) and (3.4) hold as equalities and (3.2) is satisfied by assumption.⁹ This is true in every equilibrium. Furthermore

$$\pi_H c_H - \pi_L c_L = 1 - \left[1 - (1 - \pi_L) \frac{r_L}{D} \right] = (1 - \pi_L) \frac{r_L}{D} > 0$$

that is, Ψ may be greater or lower than zero. As a result, the managers do not always invest in the same portfolio in the whole parameter space. Which credit portfolio the bank actually selects depends on the calibration of the variables. It is important to point out, however, that the private bank is efficient, in that it always chooses the portfolio with the highest NPV. This result holds

⁹It is straightforward to show that this is also true for the participation constraint when the L-portfolio is more attractive.

3 Is direct market discipline in banking really effective and necessary?

independently of the assumption that $\pi_L > \pi_H$. To see this, note that the H-portfolio has a higher value added if

$$\pi_H R_H \geq \pi_L R_L + (1 - \pi_L) r_L,$$

which is the same condition as (3.1) given the equilibrium values of c_H and c_L . On the other hand, if the bankruptcy costs are high enough, the L-portfolio may be preferred by the regulator even if it is less efficient and thus a trade-off between bank efficiency and systemic stability may emerge.

By applying a comparative static analysis it is easy to see that when the general probability of success increases (i.e., π_H and π_L go up by roughly the same amount $\Delta\pi$), the risky portfolio becomes more efficient and attractive for the bank. One can interpret some (weighted) average of π_H and π_L as a proxy for the general economic environment, since during recessions credit defaults are more frequent.¹⁰ The above implies that private banks tend to invest in the H-portfolio during phases of economic expansion and in safer assets when the economy slows down. This agrees with common opinion that credit policies of banks become more conservative during economic downturns.

Shareholder-value-maximizing public bank

Consider now a public institution whose debt is guaranteed by the state. If that bank becomes insolvent, the authorities bail it out by injecting additional capital to cover the deficit. Debt investors will never incur losses. As a result D can be borrowed at the risk-free rate. Suppose for a moment that the managers maximize the value of equity (I will return to this assumption below). Thus, the managers choose the H-portfolio if

$$\begin{aligned} \pi_H (R_H - D) &\geq \pi_L (R_L - D) \\ \text{s.t. } \pi_H (R_H - D) &\geq E \end{aligned} \tag{3.5}$$

Both conditions are always satisfied in my setup. Intuitively, given that the borrowing costs are independent of the portfolio chosen, managers care only about expected revenues, which are higher for the risky portfolio by assumption. Hence, if the L-portfolio is more efficient in NPV terms, the state guarantee may be detrimental from two points of view. First, it leads to an inefficient allocation of resources, since the H-portfolio generates a lower value added. Secondly, it raises the probability of a bank going bankrupt, as managers tend to maximize the risk of their credit portfolio. That debt guarantees have moral hazard consequences is a standard result of the literature on deposit insurance.¹¹ The assumption of a monopolistic bank is not crucial for this result. Even if the bank is perfectly competitive, it will always have the incentive to invest in the H-portfolio.

¹⁰Lehmann and Manz (2006) provide empirical evidence that the lending performance of Swiss banks, and in particular the amount of provisions, is positively correlated with the business cycle.

¹¹See, e.g., Merton (1977).

Furthermore, it is important to stress that the bank maximizes shareholder value and not profits. The allocation of resources may be inefficient since the managers – given the limited liability of equity – do not care about the recovery value in the case of default. If they maximize profits, they would always choose the most efficient portfolio. Instead, managers only look at the expected value of equity in the event of success. Note, moreover, that even by relaxing the assumption that $\pi_H R_H > \pi_L R_L$, limited liability and the absence of market discipline may induce the managers to choose the H-portfolio!

On average a shareholder-value-maximizing public bank should have higher gross interest revenues than a private institution, since it always chooses the portfolio with the maximum expected cash flow given success. On the other side, provisions are also larger, as for the H-portfolio there is no recovery and default probabilities are higher. Moreover, the private bank would have, on average, higher net interest revenues,¹² since it always invests efficiently. Finally, the business cycle ($\Delta\pi$) does not influence the choice of the portfolio and consequently the risk appetite of the managers is not procyclical as in the previous section.

Agency costs

Let me come back to the assumption that the managers of the public bank maximize shareholder value. Inequality (3.5) implicitly takes for granted that the equity-holders are protected by limited liability if the bank becomes insolvent. However, this is not true if the public bank enjoys a guarantee *and* is wholly owned by the state. In this case the state, as equity-holder, has no reason to exploit limited liability, since if the bank defaults, it will have to refund the debt holders. In other words, when maximizing shareholder value, the managers must take into consideration the losses that may be caused by the state guarantee. So how can I justify the empirical implications derived above? One could argue that the managers of public banks undergo less strict monitoring by their equity-holder when compared to managers of private banks. As a corollary, a “flippant” supervisor could be inclined to oversee public banks less accurately, since, thanks to the state warranty, their survival is guaranteed by definition. If all this is true, managers of public banks may have fewer incentives to carefully analyze the risk profile of potential borrowers, since such an effort is costly. For the sake of simplicity, suppose that this induces the managers to assign to both portfolios the same probability of default π , with $\pi_L > \pi > \pi_H$. Following the previous discussion, π can be interpreted as a proxy for the general economic environment. Then the managers choose the H-portfolio if

$$\pi (R_H - D) - (1 - \pi) D \geq \pi (R_L - D) + (1 - \pi) (r_L - D)$$

i.e., when

$$\pi R_H \geq \pi R_L + (1 - \pi) r_L \quad (3.6)$$

¹²I define net interest revenues as gross interest revenues minus provisions.

3 Is direct market discipline in banking really effective and necessary?

which is more often satisfied than in the private bank case. The empirical implications for the lending performance of the two bank groups are then the same as above, albeit the expected differences would be smaller in absolute terms.

There are of course other channels through which the more superficial monitoring could induce the managers to be more risk-seeking than the state as equity-holder. For instance, the managers may obtain private benefits by giving uncollateralized loans to relatives and friends without caring about the creditworthiness of the borrower – an H-type credit.¹³ In addition, if the compensation scheme for the executives is nonlinear, e.g., option-based, the H-portfolio may represent the more attractive investment opportunity.

In this context I am explicitly assuming the existence of agency costs between the managers and the state as unique shareholder. Several arguments can be made to justify the assumption that public banks are subject to agency costs while private banks are not. First, private shareholders have stronger incentives than some public authority to accurately monitor managers' activities, since it is their own money that is at stake and not the money of a third party (taxpayers in the case of public banks). Second, it may be easier for public bank managers to defend any particular lending policy given the often very hazily defined public bank mandate.

On the top of this, note that apart from the state, the other equity-holders of a public bank (if any) have the incentive to exploit limited liability. In Switzerland about the half of the cantonal banks are listed on a stock exchange. In such cases it is unclear how much influence private equity-holders have on the managers of those cantonal banks. If the influence is high (i.e., when the control exercised by the state is weak), the setup will be close to the one described by equation (3.5). That is, for banks that enjoy a state guarantee but are controlled by private shareholders moral hazard induced by limited liability may nonetheless affect managers' investment decisions.

Economic-value-maximizing public bank

Assume again a world with no agency costs between managers and the state. We saw in the previous section that it is somewhat unreasonable to assume that the managers of a public bank exploit the limited liability status of equity. So do public institutions maximize profits? Although this is a standard assumption in the literature on industrial organization, state-guaranteed institutions may have different objectives. For example, the public mandate explicitly requires most Swiss cantonal banks to support the local economy. One way to interpret this mandate is that the bank should lend money to all projects with positive NPV, or, if some rationing is needed, to the projects with the greatest NPV. Given the assumptions of the model, this is equivalent to profit-maximizing. In a static world with market discipline both bank groups would then finance the same firms. This is consistent with a general equilibrium model with no frictions, where a shareholder-value-maximizing firm behaves like a welfare-maximizing institution.

¹³This kind of misuse of public banks can arise even in the absence of agency costs between managers and the government. For instance, the government may try to control public banks in order to pursue political rather than social objectives (see La Porta, Lopez-De-Silanes, and Shleifer, 2002).

Public bank sets price equal to marginal cost

In the previous sections I assumed that each bank has the bargaining power in arranging the terms of a credit and sets the price so as to maximize interest income. As a result, borrowers were indifferent as to whether to go to a private bank or a public bank and the loan interest rate had a markup over marginal costs. In order to support the local economy, public banks may set this markup equal to (or near) zero. In a model with the risk-free rate normalized to one, the public bank has a zero marginal cost and therefore sets a price equal to $\frac{1}{\pi_H}$ for the H-portfolio and equal to $\frac{1-(1-\pi_L)r_L}{\pi_L}$ for the L-portfolio. The expected net interest income is zero. Obviously, such a pricing policy makes no sense if the public institution is a shareholder-value- or profit-maximizer. But it is compatible with other objectives, such as the maximization of economic value added. Note that in this case every borrower would like to go to the public bank, but because of the limited amount of funds available just a few “lucky” firms will be financed. Moreover, a zero-markup pricing policy does not influence the portfolio choice of the public bank, so only efficient borrowers are supported by both institutions. And since the public bank sets a lower markup, it should have lower gross and net interest revenues.

In practice, marginal costs are strictly greater than zero even for public banks. However, Bichsel (2006) and Facchinetti (2007) showed that cantonal banks have lower funding costs than the typical private bank in Switzerland.

To sum up, the empirical implications for the portfolio choice and performance of public banks strongly depend on their utility function. Given the typical mandate – to support the local economy and local business enterprises – one should expect public banks to maximize economic value added. At the same time, however, they should also generate enough profits: no politician is likely to put up with an underperforming bank. As discussed above, in this perfect world with market discipline and no frictions in the public sector, all banks should behave very similarly.

3.2.2 No market discipline

Suppose now that the market does not discipline private banks through a risk-sensitive supply of debt. The H-portfolio would then represent an optimal investment if

$$\begin{aligned} \pi_H (R_H - cD) &\geq \pi_L (R_L - cD) \\ \text{s.t. } \pi_H (R_H - cD) &\geq E \end{aligned} \tag{3.7}$$

where c is the risk-insensitive marginal cost of debt. Equation (3.7) reduces to

$$\pi_H R_H - \pi_L R_L + (\pi_L - \pi_H) cD \geq 0.$$

Since $\pi_H R_H > \pi_L R_L$ and $\pi_L > \pi_H$ by assumption, in the static game private banks always choose the H-portfolio as the shareholder-value-oriented public bank of Section 3.2.1. Indeed, when $c = 1$

3 Is direct market discipline in banking really effective and necessary?

both types of bank will face exactly the same problem. Therefore, if the public institution is an economic-value-maximizer but does not set the price equal to marginal cost, it should have higher average net interest revenues but lower provisions and gross interest revenues, since sometimes the managers will invest in the safer and more efficient credit portfolio. Given (3.6), the same is true when there are agency costs between managers and state.

When $c > 1$, i.e., when debt investors are uninformed but require a flat risk premium on the money they lend, the incentives of the private bank to choose the H-portfolio are even greater than those of the shareholder-value-oriented public bank. This result is in line with Stiglitz and Weiss (1981), who showed that in credit markets with imperfect information higher interest rates induce borrowers to undertake riskier projects. Intuitively, since equity can be seen as a call option on bank assets, a higher strike price (the amount of debt at maturity including interests) increases the relative importance of the limited liability effect, making the H-portfolio more attractive. In Appendix 3.7.2 it is shown that this is also true in a more realistic Black-Scholes world.

The simple model above has several implications for the lending business of public and private banks that can be tested. Table 3.2 summarizes the findings in a static context.

Table 3.2: Empirical implications of the static model

“Goal” of the public bank	Market discipline on private banks	
	Effective	Ineffective
Maximize shareholder value (exploit limited liability)	I: Higher [†] P: Higher NI: Lower	I: Equal P: Equal NI: Equal
Superficial monitoring (e.g., $\pi_L = \pi_H = \pi$)	I: Higher P: Higher NI: Lower	I: Lower P: Lower NI: Higher
Maximize economic value with markup pricing	I: Equal P: Equal NI: Equal	I: Lower P: Lower NI: Higher
Maximize economic value with $P=MC$	I: Lower P: Equal NI: Lower	I: Lower P: Lower NI: ?

I: Interest revenues; P: Amount of provisions; NI: Net interest revenues (=I-P)

[†]For a public bank compared to a private bank. In other words, the static model implies that in the presence of market discipline the shareholder-value-maximizing public bank generates, on average, higher interest revenues than the private bank.

3.3 Some remarks in a dynamic framework

In a dynamic game things are a little different. Several studies¹⁴ show that the charter value, defined as the future intrinsic value of the bank as an ongoing institution, may significantly affect

¹⁴See the introduction.

3.3 Some remarks in a dynamic framework

the risk choice of the management. In the context of Section 3.2, suppose that each year the bank must decide which project it wants to finance. When the private bank invests in the riskier portfolio, it must now take into account a higher probability of losing its charter value F^* . That is, the managers invest in the H-portfolio when

$$\begin{aligned} \pi_H (R_H - c_H D + F^*) &\geq \pi_L (R_L - c_L D + F^*) \\ \text{s.t. } \pi_H (R_H - c_H D + F^*) &\geq E + F^* \end{aligned}$$

under the assumption that the market can discipline banks. For the banks in my model, F^* is the present value of the expected net income stream, given an optimal lending policy in the future.¹⁵ I assume that debt has a repayment period of one year, so that c_H and c_L are the same as before. The first condition above reduces to

$$\pi_H R_H - \pi_L R_L - (\pi_H c_H - \pi_L c_L) D - (\pi_L - \pi_H) F^* \equiv \tilde{\Psi} \geq 0.$$

Since $\tilde{\Psi} < \Psi$, in a dynamic context the private institution is more conservative as it chooses the L-portfolio (if any) more often than in the static case. Intuitively, if the franchise value of the bank is high, its managers and shareholders do not want to lose it and prefer to invest in relatively safe assets. This is true when

$$F^* \geq \frac{\pi_H R_H - \pi_L R_L - (1 - \pi_L) r_L}{\pi_L - \pi_H}. \quad (3.8)$$

Given the results in the static framework, this lending policy may be inefficient from a general economic perspective, especially if there are no externalities associated with banks' defaults. The H-portfolio may have a higher NPV, but the managers would still choose the L-portfolio when they are not compensated for the higher risk of losing F^* . If (3.8) is satisfied in the whole parameter space, the private bank always minimizes risks. In addition, note that during economic downturns (π_L and π_H shrink by roughly the same amount) the managers face a trade-off. On the one hand it seems reasonable to assume that the franchise value of a bank is positively correlated with the cycle, that is $\frac{\partial F^*}{\partial \pi} > 0$. As a result, the bank's managers may be inclined to assume more risks when the economy performs poorly. On the other hand, the safer investment becomes more efficient and attractive for the bank, as in the static case. The net effect may therefore be ambiguous. This is consistent with Saunders and Wilson (2001), who provide evidence that the relationship between charter value and bank risk-taking incentives is sensitive to market conditions.

If market discipline is ineffective, the private bank invests in the H-portfolio when

¹⁵If one assumes that at the beginning of each period the bank always has the same investment opportunities, F^* can be exactly calculated, although with $r_f = 1$ we have $F^* \rightarrow \infty$. In reality, however, F^* is a random variable that depends on several factors. To keep things simple, I only assume that $F^* > 0$.

3 Is direct market discipline in banking really effective and necessary?

$$\begin{aligned} \pi_H (R_H - cD + F^*) &\geq \pi_L (R_L - cD + F^*) \\ \text{s.t. } \pi_H (R_H - cD + F^*) &\geq E + F^*. \end{aligned}$$

As before, one can rewrite the first condition as

$$\pi_H R_H - \pi_L R_L - (\pi_H - \pi_L) cD - (\pi_L - \pi_H) F^* \equiv \hat{\Psi} \geq 0.$$

Again, which portfolio the managers actually choose strongly depends on the value of F^* . If the franchise value is high enough, the private bank will be cautious in its credit policy. In addition, note that since $(\pi_H - \pi_L) c < 0$ while $\pi_H c_H - \pi_L c_L > 0$, we have $\hat{\Psi} > \tilde{\Psi}$, and therefore also in a dynamic context market discipline tends to reduce the risk appetite of the managers of a private bank. However, this is true only if one keeps F^* constant. All else being equal, a bank with low debt costs (low c relative to c_H and c_L) would also have a larger franchise value. As a result, it is a priori unclear whether market discipline increases or reduces the risk appetite of bank managers. But even if market discipline decreases portfolio risk, it may not be socially desirable. We have seen that disciplined banks with high F^* may invest too conservatively. On the other hand if c is large enough, a risk-insensitive supply of funds may increase welfare. As a corollary, the model implies that market discipline is likely to be more effective and appealing when the expected franchise value is low, for instance in periods of stress for the banking business.

The empirical implications of the dynamic model for the lending business are less clear cut, since the value of F^* plays a key role. If the franchise value of the bank is low, the previous results apply. If it is large, however, whether market discipline is effective or not becomes irrelevant.

The behavior of the fully public institution is generally not influenced by F^* . In fact, the probability that the equity-holder loses the franchise value is zero since, by assumption, the bank will always be bailed out in the case of default. If, however, the public bank is currently controlled by private shareholders who would lose their stake in the event of a bailing-out, this may not be true. F^* would then influence the investment decision process in a similar manner as in the private bank case. However, historical evidence in Switzerland suggests that private shareholders generally do not lose their claims when a cantonal bank is recapitalized by the state.¹⁶ The above implies that the state guarantee may prevent high-charter-value banks from being excessively risk-averse and thus possibly generate economic value added, which would justify the existence of public banks even in developed countries.

¹⁶For instance, the recapitalization of Banque Cantonale Vaudoise (BCV) in 2002 was carried out by the Canton of Vaud.

3.4 Empirical analysis

3.4.1 Methodology and objectives

The model above offers a good starting point for an empirical analysis of whether the market can directly discipline Swiss private banks. One basic assumption of the model is that private banks' lenders monitor the risk profile of each bank and invest accordingly. There is enough evidence that capital market investors – bondholders in particular – “punish” banks if they assume excessive risks. Yet, the great majority of Swiss banks do not raise funds from the capital markets, the typical channel through which market discipline is expected to work. However, Birchler and Maechler (2002) provide empirical evidence that (regional) bank depositors – the most important lenders of the typical Swiss bank – also react to bank-specific fundamentals. In particular, they find that “the perceived ‘quality’, or relative safety, of a bank clearly influences depositors’ willingness to hold uninsured savings deposits in a particular bank”.

In order to test if such a risk-sensitive supply of funds influences the risk appetite of private banks' managers, I compare the lending performance – i.e., interest revenues and P&L provisions – of three domestic bank groups: cantonal banks, regional banks and, to a lesser extent, the Raiffeisen group. All the institutions included in the sample are primarily active in the Swiss retail banking market and therefore have similar characteristics. The main difference lies in the fact that cantonal banks benefit from a state guarantee¹⁷ and, as a corollary, may not maximize shareholder value but pursue other objectives. This may also be true for the Raiffeisen group, since its organization is based on cooperative principles. The two Swiss big banks, UBS and Credit Suisse Group (CSG), are not included in the sample, since – given their size, sophistication and internationalization – UBS and CSG differ substantially from the other Swiss banks. The empirical implications of the model derived in Section 3.2 and Section 3.3 can be directly tested by comparing the lending performance of regional (private) and cantonal (public) banks.

The analysis of lending performance before interest expenses instead of other measures of risk or performance has several advantages.

First, it allows us to disentangle the effects the state guarantee and public ownership have on risk-taking from the effects they have on funding, cost efficiency and other banks' characteristics. This should make the empirical results more robust, since it is easier to formulate an empirical model for lending performance than for other more comprehensive measures, especially for those banks that are involved in more banking activities.

Second, the testable implications I derived from the theoretical model directly concern the asset side of the credit business. Running estimations on other variables such as leverage, interest rates offered or total profitability without the backing of a theoretical model is dangerous and may lead to misspecifications and misinterpretations of the results. The focus on lending performance, both

¹⁷For Banque Cantonale Vaudoise (BCV) the guarantee is implicit, while for Banque Cantonale de Genève (BCGe) it is limited. All other cantonal banks enjoy an explicit, full state guarantee (although in some cases the guarantee does not cover subordinated debt).

3 Is direct market discipline in banking really effective and necessary?

in the theoretical as well as in the empirical section, differentiates this study from the existing literature. As mentioned before, all banks in the sample face the same problem: lend money so as to maximize their specific utility function. This screening (and monitoring) function is one of the basic rationales that justify the mere existence of banks and direct market discipline can be seen as an instrument that optimizes the incentives in this perspective. As lending performance measures the efficiency of screening and monitoring, it represents, in my view, a simple but effective basis for testing direct market discipline. As a corollary, it is relatively straightforward to derive welfare implications by comparing lending performance across banks.

Finally, the available data set on interest revenues and provisions is quite accurate when compared with other risk measures, loan portfolio composition in particular.

3.4.2 Model specification

The following three equations for gross interest revenues $I_{i,t}$, P&L provisions $P_{i,t}$ and net interest revenues $NI_{i,t}$ as a proportion of total loans define my empirical model:

$$\text{Gross interest revenues}_{i,t} = f_I(X_{i,t}^I, \beta^I, v_{i,t}^I) \quad (3.9)$$

$$\text{Provisions}_{i,t} = f_P(X_{i,t}^P, \beta^P, v_{i,t}^P) \quad (3.10)$$

$$\text{Net interest revenues}_{i,t} = f_{NI}(X_{i,t}^{NI}, \beta^{NI}, v_{i,t}^{NI}). \quad (3.11)$$

$X_{i,t}^I$ and $X_{i,t}^P$ are vectors of explanatory variables, β^I , β^P and β^{NI} vectors of parameters. $X_{i,t}^{NI} = X_{i,t}^I \cup X_{i,t}^P$ for every bank i and year t . $\nu = [v_{i,t}^I \ v_{i,t}^P \ v_{i,t}^{NI}]'$ is the vector of the error terms. I assume that $\nu \sim d(0, \Omega)$, where Ω is non-diagonal: Some unexpected shocks affect both gross interest revenues and provisions and $v_{i,t}^{NI}$ is a function of $v_{i,t}^I$ and $v_{i,t}^P$ by definition. In addition, the errors may be autocorrelated for at least two (related) reasons. First, it is difficult to fully restructure a credit portfolio within one year, especially for small banks that are not active in the credit derivative and securitization market. As a result, a portfolio which is relatively more profitable today, is also likely to be relatively more profitable tomorrow. And secondly, I argue below that ν may have an unobservable component – the risk appetite of bank i – which is to some extent persistent over time. Moreover, the errors are likely to be heteroskedastic since some banks are better diversified than others and hence less exposed to fluctuations in earnings. For example, Ersparnisgesellschaft Küttigen, a regional bank in the sample, has just one local office and 80 per cent of its assets invested in mortgages. On the other hand, Zürcher Kantonalbank (ZKB) manages more than hundred branches and is active in almost every credit and mortgage market segment. I tested for serial correlation and heteroskedasticity using the Wooldridge (2002) test for serial correlation and an LR test in the idiosyncratic errors of my basic linear panel-data model (see Section 3.5.1). The null hypothesis of no first-order autocorrelation can be rejected at the 1 per cent significance level, with the exception of $v_{i,t}^P$ (rejection at 5 per cent significance). Homoskedasticity is rejected at the 1 per cent level in all cases.

The explanatory variables employed in the estimations, i.e., the elements of $X_{i,t}^I$ and $X_{i,t}^P$, are described in Table 3.3. With the exception of CB dummy_i, they are all control variables.

Table 3.3: Description of the explanatory variables

Name	Description
GDPgrowth _t	Annual growth rate in real Swiss GDP.
Libor _t	Annual average 3-month LIBOR rate (BBA) denominated in Swiss francs.
YieldCurve10Y _t	Difference between the average 3-month LIBOR rate and the average 10-year Swiss Confederation bond yield.
Default _t	Total amount of domestic corporate and private defaults to total loans.
REcrisis _{i,t}	1 if bank <i>i</i> faced a real estate crisis in year <i>t</i> ; 0 otherwise. A crisis occurs when: (i) the weighted real estate price index for bank <i>i</i> falls in <i>t</i> and (ii) in either <i>t</i> -1, <i>t</i> -2 or <i>t</i> -3 it increased by more than 2 per cent.
Area _{i,t}	Sum of the areas in thousand km ² of 8 pre-defined Swiss regions, weighted by the exposure of bank <i>i</i> in each region.
Pop _{i,t}	Sum of the population (in 2001) in each region, weighted by the exposure of bank <i>i</i> in that region.
BankNr _{i,t}	Sum of the total number of loan and mortgage banks active in each region, weighted by the exposure of bank <i>i</i> in that region.
CB dummy _i	1 if cantonal bank; 0 otherwise.
Raiffeisen dummy _i	1 if Raiffeisen group; 0 otherwise.
DataBreak dummy _t	1 date of observation prior to 1995; 0 otherwise

The regressors can be grouped in three classes, depending on the reason for their influence on lending performance. All of them may affect both gross interest revenues and provisions so that $X_{i,t}^I = X_{i,t}^P$.

General operating conditions

The essence of financial intermediation is to borrow money at a given cost and lend the funds raised at a slightly higher price. The main determinant of the borrowing costs is the risk-free rate. The higher this rate, the more gross interest revenues the bank has to generate to cover its expenses. In general, every change in the yield curve affects the pricing policy of banks and their interest revenues. By including the term structure in the estimation (YieldCurve10Y_t), I am also able to capture changes in expectations. Expectations are particularly important, since in many credit contracts the interest is set for a relative long period of time – e.g., fixed rate mortgages may have a term of several years. In addition to this direct effect, a rise in the risk-free rate has, *ceteris paribus*, a positive effect on provisions as well, as banks' borrowers would face increasing

3 Is direct market discipline in banking really effective and necessary?

debt costs, which could lead to more insolvencies. Note that such an indirect interest rate risk also negatively influences interest revenues (some payments are omitted), although the positive effect described before should be dominant. To better capture the effect of borrowers' creditworthiness on lending performance, I include the number of corporate and household defaults as an explanatory variable.

Another key driving factor of the lending performance of a bank is the business cycle. During periods of economic expansion, the demand for credits increases, which should allow the banks to extract a higher rent from the money they lend.

Finally, in the mortgage business especially, it is important to control for the value of the collateral. In the early nineties, a real estate crisis plunged many Swiss banks into difficulties, forcing them to build up huge provisions. Real estate price indices represent a good proxy for the quality of mortgages. Since there may be substantial regional differences in the evolution of real estate prices, I elaborated a weighted index. I partitioned the Swiss market into 8 regions.¹⁸ The relevance of each region for every bank was then determined by looking at the percentage of the mortgages in that region relative to the total mortgages of the bank. Finally, those weights were multiplied with a regional real estate price index and summed up. The real estate price indices come from Wüest & Partner, a consultancy firm for real estate in Switzerland. In general, the collateral becomes important when the borrower goes bankrupt. Moreover, in such circumstances banks are probably more affected by the downside potential of the collateral value. Hence, the real estate price index is supposed to have a nonlinear effect on interest revenues and provisions. In order to control for this, the variable I use in the regressions ($REcrisis_{i,t}$) is a dummy that equals one if bank i faced a real estate crisis in year t . This occurred when the institution-specific real price index defined above fell in t and increased by more than 2 per cent in at least one of the three previous years.

Clearly, all the variables described above are strongly interdependent, so that collinearity problems could emerge, making the interpretation of the coefficients problematic. Yet, since they are all control variables, I am less interested in the interpretation of their effects on lending performance.

Level of competition

A priori in a less competitive environment a bank can extract a higher rent from its financial intermediation business. Even in a small country like Switzerland there may be significant differences across regions and segments with respect to the degree of competition among banks. Several proxies are used to control for the effect of competition on lending performance. To construct those indicators I employed a method similar to the one described in the previous section. In order to ascertain the importance of each region for a particular bank I focused on the percentage of the branches in that region relative to the total branches of the bank. The weights were then multiplied with several proxies for the level of competition in each region: the number of active banks, the population and the area of that region. I assume that competition is tougher if, ceteris

¹⁸Zurich region, eastern Switzerland, central Switzerland, north-western Switzerland, Berne region, southern Switzerland, Lake Geneva region and western Switzerland.

paribus, more banks are active in a particular market, if the demand for credit is low and if the “transportation” costs are low. Also those variables may be collinear, though. For instance, when a densely populated region is particularly attractive and profitable for the lending business, it is likely that a lot of banks have settled there. Again, this may complicate the interpretation of the specific coefficients, but this is not a problem in relation to the aims of my study.

Risk and market discipline

A further basic determinant of the lending performance, in addition to the more or less exogenous variables described before, is the risk profile of a bank’s loan portfolio. While systematic risk is presumably captured by the variables described in the previous sections, one also needs to measure idiosyncratic bank risk. Unfortunately this institution-specific variable depends on several factors, some of which – such as the risk appetite of the managers – are only partially observable and quantifiable. Moreover, even observable bank-specific variables that proxy idiosyncratic asset risk can be hardly used in the estimations. To illustrate the problem, suppose that

$$v_{i,t}^I = \theta_{i,t} + \xi_{i,t}^I$$

where $\theta_{i,t}$ represents the risk appetite of bank i and $\xi_{i,t}^I$ is the residual error. Let us take the ratio of mortgages to total credits ($x_{i,t}^{I,j}$) as a proxy for asset risk, as mortgages are believed to be less risky (and less profitable) than other credits. The problem is that this ratio too is likely to depend on the bank’s risk appetite, as the latter has a direct impact on the lending performance. Intuitively, a risk-seeking bank prefers to invest in unsecured, non-investment grade loans than in mortgages. In more technical terms

$$x_{i,t}^{I,j} = g(\theta_{i,t}, \dots)$$

and hence

$$E(X_{i,t}^I v_{i,t}^I) \neq 0$$

if the simultaneous effect of $\theta_{i,t}$ on interest revenues and $x_{i,t}^{I,j}$ is not perfectly captured by $X_{i,t}^I$, which is very unlikely. This issue does not concern only the ratio of mortgages to total credits, but the majority of the bank-specific variables. A basic assumption of every standard estimation model would then be violated. Note that this endogeneity problem cannot be solved by using lags (or leads) as instruments, since it is reasonable to assume that a component of $\theta_{i,t}$ is constant over time. That is, if the management of a bank is relatively risk-averse today, it will probably be risk-averse tomorrow as well.

However, not all the factors affecting $\theta_{i,t}$ are unobservable. In Section 3.2, I showed that the degree of market discipline may be an important determinant of the risk appetite of banks. Since the aim of this paper is to establish whether direct market discipline is effective or not, one somehow has to find a way to measure it. The structure of the Swiss banking sector is very helpful in this

3 Is direct market discipline in banking really effective and necessary?

sense, as an important group of financial intermediaries active in Switzerland are cantonal banks, which benefit from a state warranty. Safety nets wipe out, or at least weaken, market discipline. In order to reduce moral hazard, certain cantons demand a kind of risk-adjusted reward for the state guarantee. This compensation is usually based on the amount of required capital held at year end.¹⁹ However, the discipline exercised by the market is expected to be stronger, since it is continuous and not limited to a specific aspect of the bank's risk profile. My idea is therefore to compare the lending performance of cantonal banks with the performance of regional banks. Regional banks represent a good benchmark as they are predominantly active in the same business as cantonal banks but are privately owned and hence possibly subject to direct market discipline. A further difference lies in the size: at the end of 2003, total credits outstanding at the average cantonal bank amounted to 10 billion Swiss francs, while the figure for the average regional bank was slightly more than 0.8 billion. Hence, if positive economies of scale exist in lending, cantonal banks should be, *ceteris paribus*, more profitable.

Technically, I introduce a dummy variable in the estimation which equals 1 if the bank enjoys a state guarantee (implicit or explicit) and 0 otherwise. Note that although market discipline should affect the risk appetite of a bank, that bank does not have the possibility of deciding whether it can enjoy a state guarantee or not. From this perspective, the dummy is exogenous. On the other hand, one can also argue that different compensation structures, the absence of market discipline or less stringent monitoring by the state as equity-holder may attract particularly risk-loving managers. This adverse selection problem would make the dummy endogenous. Yet, there is little evidence that performance and risk play a more important role as determinants of executive compensations at cantonal banks. Moreover, risk-averse managers too would profit from the absence of market discipline or less stringent monitoring by the equity-holder. In other words, every manager likes to go to a public bank if such frictions exist. I tested the exogeneity assumption of CB dummy_i with a Hausman (1978) specification test and could not reject the null.²⁰

A further issue concerns the role of the franchise value, since it is reasonable to assume that in reality banks optimize their investment decisions over the long run. As shown in Section 3.3, the value F^* represents a key determinant of a bank's portfolio choice and, consequently, of its lending performance. The franchise value is also hardly quantifiable, though, as it depends on unobservable bank characteristics such as the value of relationships and human capital. In related studies that focus on publicly traded institutions, F^* is often proxied by an estimate of banks' Tobin's q (see, for example, Keeley, 1990 or Demsetz, Saidenberg, and Strahan, 1996). Unfortunately, as the majority of the banks included in the sample are not listed on a stock exchange, I have to rely on other

¹⁹The law governing Glarus Cantonal Bank (Art. 5), for instance, prescribes that the annual payment to the canton should amount to 0.3% of required capital under the Swiss Banking Act. Cutoff date is 31 December. Other cantons have similar rules. This may explain why the required capital of cantonal banks tends to decrease every year between September and December.

²⁰Similarly, I tested the hypothesis that balance sheet variables are exogenous but I had to reject the null. In all cases, however, it was extremely difficult to obtain a positive definite variance matrix of the difference between the estimators ($\beta^{consistent}$ and $\beta^{efficient}$ respectively) and so these results should be treated with some caution.

methods to control for the franchise value. These are described below.

A final point is worth noting: As I do not know what goals the cantonal banks actually pursue, it may be difficult to identify unambiguously the reason for possible differences in lending performance. For instance, suppose that $f_I()$ and $f_P()$ are linear and public banks maximize shareholder value, including the value of limited liability. Then my static model implies that, in the presence of market discipline, the intercepts β_0^I and β_0^P in a subsample of cantonal banks are expected to be higher than the same intercepts in a subsample of regional banks, while β_0^{NI} should be lower. Yet, even if all the intercepts are equal one cannot conclude that market discipline is ineffective. In fact, if cantonal banks maximize economic value added with markup pricing then, in the presence of market discipline, the lending performance of the two bank groups will be the same.

3.4.3 Data set

The data set consists of an unbalanced panel of 29 cantonal and 214 regional banks. A time series for the Raiffeisen group is also available, but not for each single Raiffeisen bank. Observations were collected on a yearly basis from 1987 until 2004 and come from the internal database of the Swiss National Bank (SNB). Every bank uses the same accounting standards defined by the Swiss Federal Banking Commissions (SFBC)²¹ when reporting to the SNB. Thus, comparisons across institutions and banking groups are not biased by differences in accounting rules. On the other hand, there may be a structural break in the data as new reporting principles were introduced between 1995 and 1996 by the SFBC and the SNB. In particular, until 1995 banks were not required to separately disclose provisions for credit risk and depreciation on non-current assets. For that reason I introduced DataBreak_{it} in my model. In general, this change in rules does not appear to significantly affect the main results.

All banks in the sample are mainly active in the Swiss mortgage and loan business. Only a few of the biggest institutions engage significantly in banking activities other than traditional financial intermediation. Moreover, the great majority of regional banks do not raise funds through the capital markets. The Raiffeisen group has a federal structure and is based on cooperative principles. As a result, each member of the group may not maximize shareholder value but pursue other objectives. The representative Raiffeisen bank is a very small institution mainly active in the mortgage market. Raiffeisen banks are present in almost every Swiss region.²² Table 3.4 summarizes the main characteristics of the sample.

Interestingly, the cantonal banks have higher interest revenues on average than regional banks but also usually need to build up a larger amount of provisions. However, the difference of the means is statistically significant only for the provisions. This would suggest that public banks tend to invest in riskier projects than regional institutions, but those projects are not more profitable on average.

²¹See http://www.ebk.admin.ch/d/archiv/2002/rrv-annexe/rrv_d.02.pdf for a description of the guidelines in German. Unfortunately there is no official English translation.

²²See Rime and Stroh (2003) for a more detailed description of the structure of the Swiss banking sector.

3 Is direct market discipline in banking really effective and necessary?

Table 3.4: Summary statistics

All ratios in percent.						
Variable	Cantonal		Regional		Raiffeisen [†]	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Interest revenues to total loans	5.49	1.40	5.43	1.25	5.25	1.33
Provisions to total loans	0.62	0.70	0.52	0.46	0.35	0.20
Net interest revenues to total loans	4.87	1.39	4.91	1.17	4.90	1.25
Total loans (in bn CHF)	7.76	8.79	0.48	1.08	47.93	21.08
Mortgages to total loans	67.08	26.20	78.91	19.86	85.13	5.00
First rank mortgages [‡] to total mortgages	92.31	6.63	93.67	6.62	94.58	1.39
Total assets (in bn CHF)	9.93	12.05	0.57	1.24	57.23	25.51
Return on assets (RoA)	0.27	0.35	0.31	0.25	0.21	0.15
Return on equity (RoE)	7.42	9.00	6.39	2.87	7.75	2.11
Leverage	95.79	0.89	94.66	2.65	96.90	1.07
Domestic branches	44	38.74	5	8.77	1298	32.34
Number of observations	468		2506		18	

[†]Data at the group level. [‡]First rank mortgages are mortgages with a low loan-to-value (LTV) ratio.

As expected, cantonal banks are, on average, bigger than regional banks and usually have more branches scattered across the territory. The statistically significant differences in the ratio of mortgages to total loans, profitability or leverage corroborate the hypothesis that this kind of balance sheet information may be endogenous in my setup.

The Raiffeisen group appears to be even more conservative than regional banks – possibly because Raiffeisen banks are active almost exclusively in the high-quality mortgage business – but no less profitable: average net interest revenues are almost the same for all bank groups. Unfortunately, since I have data only at the group level, it makes no sense to test whether these differences are significant.

It is important to point out that the aforementioned real estate crisis in the early nineties gave rise to a wave of consolidation in the Swiss banking sector. Several ailing regional and cantonal banks were acquired by larger institutions (mainly the two big banks) and hence dropped out of the sample. For instance, between 1990 and 1996 the number of cantonal banks fell from 29 to 24, while the number of regional banks plunged from 203 to 119. Hence, the sample is historically unbalanced and the averages above may therefore be biased. Below, therefore, I control for exogenous factors that may have affected the performance of the banks over time.

3.5 Results

The exact functional form of $f_I()$ and $f_P()$ is not known. Nevertheless, I use a parametric approach in estimating my econometric model: I make simplifying assumptions about the functional form of $f_I()$ and $f_P()$. For simplicity's sake, I opt for a transcendental logarithmic (translog) specification. This choice is mainly motivated by the size of my data sample, which, in my opinion, does not allow an efficient estimation of more complex models. The translog function is a fairly common specification in this branch of the literature.²³ The equations (3.9) to (3.11) then reduce to

$$\ln i_{i,t} = \sum_{j=1}^J \beta_j^I \ln y_{i,t}^{I,j} + \sum_{j=1}^J \sum_{z=1}^Z \gamma_{j,z}^I \ln y_{i,t}^{I,j} \ln y_{i,t}^{I,z} + \varepsilon_{i,t}^I \quad (3.12)$$

$$\ln p_{i,t} = \sum_{j=1}^J \beta_j^P \ln y_{i,t}^{P,j} + \sum_{j=1}^J \sum_{z=1}^Z \gamma_{j,z}^P \ln y_{i,t}^{P,j} \ln y_{i,t}^{P,z} + \varepsilon_{i,t}^P \quad (3.13)$$

$$\ln ni_{i,t} = \sum_{j=1}^J \beta_j^{NI} \ln y_{i,t}^{NI,j} + \sum_{j=1}^J \sum_{z=1}^Z \gamma_{j,z}^{NI} \ln y_{i,t}^{NI,j} \ln y_{i,t}^{NI,z} + \varepsilon_{i,t}^{NI} \quad (3.14)$$

where all the variables, with the exception of the dummies, were transformed to make them translog-compatible. For instance $p_{i,t} = 1 + \frac{\text{Provisions}_{i,t}}{\text{Total loans}_{i,t}}$ and $y_{i,t}^{I,2} = 1 + \text{Libor}_t$. For all j and z I can impose the symmetry restrictions $\gamma_{j,z} = \gamma_{z,j}$ without loss of generality.

²³See, for instance, Rime and Stiroh (2003).

3.5.1 Cobb-Douglas model

To identify a general effect of the cantonal bank status on lending performance, I first estimate equations (3.12) to (3.14) without all interaction terms. In other words, I approximate the translog specification with a more restrictive Cobb-Douglas-type model. The main results are shown in Table 3.5.²⁴ Given the special form of the variance-covariance matrix discussed earlier, I employ a feasible generalized least squares (FGLS) estimator for cross-sectional time-series. This approach allows for the presence of panel-specific AR(1) autocorrelation and heteroskedasticity across panels and will be used throughout the paper.

The results from samples A and B suggest that cantonal banks tend to assume more risks than regional banks and generate higher gross interest revenues, as the coefficients of the cantonal bank dummy in the respective equation are significantly greater than zero. For instance, the ratio of provisions to total loans of the average bank is 20 basis points or 40 per cent higher when the representative bank is a cantonal bank.²⁵ However, those additional risks are not adequately rewarded: In terms of net interest revenues regional banks significantly outperform cantonal banks. A surplus of 11 basis points may look small at first glance given average net interest revenues to total loans of about 5 per cent. Yet, after financing costs – the LIBOR rate in the sample averages 4.4 per cent – the difference in performance becomes economically important. Conversely, between the Raiffeisen group and regional banks there are no significant differences in the lending performance. According to my static model, one explanation for these results is that regional and cantonal banks maximize the value of equity and that market discipline prevents regional banks from assuming excessive risks. Alternatively, one can argue that the managers of public banks are subject to less strict monitoring by the owner and hence are less careful in lending money. Of course, both explanations may be true at the same time. In any case, these results provide evidence that the market exercises some direct discipline on private banks. In fact, the static model predicts that without market discipline the private bank would never be less aggressive than the public bank.

As I only have detailed data on corporate and private defaults back to 1993, I estimated the Cobb-Douglas model using the lagged value of the GDP growth rate instead of Default_t . Since the number of defaults often lag the cycle – indeed the correlation between Default_t and GDPgrowth_{t-1} equals -0.85 – the effect of this change on the empirical model should be minimal. On the other side, this allows me to almost double the sample size (2,979 observations). Interestingly, I find that

²⁴The whole set of results can be found in Appendix 3.7.1.

²⁵This figure (and those that follow) reflects the effect of the dummy on the average bank in the sample, i.e.,

$$\text{Mean} \left(e^{\widehat{\ln p_{i,t}}} - 1 \mid \text{CB dummy}_i=1 \right) - \text{Mean} \left(e^{\widehat{\ln p_{i,t}}} - 1 \mid \text{CB dummy}_i=0 \right).$$

Alternatively one can calculate

$$\text{Mean} \left(e^{\widehat{\ln p_{i,t}}} - 1 \mid \text{cantonal bank} \right) - \text{Mean} \left(e^{\widehat{\ln p_{i,t}}} - 1 \mid \text{regional bank} \right).$$

that is, the difference between the re-transformed mean fitted value of the dependent variable in the subsample of cantonal banks from the mean fitted value in the subsample of regional banks. The two methods produce almost the same results.

Table 3.5: The Cobb-Douglas model

The translog-transformed values of interest revenues, provisions and net interest revenues to total loans are regressed on the cantonal bank dummy and a set of transformed control variables using a FGLS model with panel-specific autocorrelations. In sample B the Raiffeisen group dummy is added. Samples C and D comprise cantonal banks as well as regional banks with a low and a high charter value respectively. All coefficients and standard errors are in basis points.

Sample of banks:	(A) Cantonal and regional banks	(B) All banks	(C) Regional banks with low charter value	(D) Regional banks with high charter value
Variable	Coefficient (Std. Err.)	Coefficient (Std. Err.)	Coefficient (Std. Err.)	Coefficient (Std. Err.)
<i>Interest revenues:</i>				
CB dummy _i	6.50* (2.93)	6.51* (2.93)	8.01** (3.02)	6.51* (3.08)
Raiffeisen dummy _i	—	-11.06 (10.88)	—	—
<i>Provisions:</i>				
CB dummy _i	20.31** (1.42)	20.40** (1.40)	17.38** (1.78)	18.14** (2.59)
Raiffeisen dummy _i	—	-12.04 (10.20)	—	—
<i>Net interest revenues:</i>				
CB dummy _i	-11.37** (2.52)	-11.29** (2.52)	-9.06** (2.68)	-14.67** (2.67)
Raiffeisen dummy _i	—	3.82 (4.34)	—	—
Number of obs.	1583	1595	845	1033
Number of banks	161	162	85	103
Significance levels :	†=10% *=5% **=1%			

3 Is direct market discipline in banking really effective and necessary?

the cantonal bank dummy is significantly positive in all three equations. That is, while cantonal banks assume more risks than regional banks, they are more profitable also in terms of net interest revenues. In contrast to my previous findings, therefore, cantonal banks no longer appear to be inefficient. I have several explanations for this result. First, between 1987 and 1993 the number of regional banks fell from 213 to 154. Presumably those regional banks that dropped out of the sample were institutions that performed relatively poorly, thus shifting the result in favor of the cantonal banks. Indeed, if one limits the sample only to those institutions that survived the whole period (balanced panel), β_{CB}^{NI} becomes insignificant. Using a balanced panel may, however, distort the analysis, since underperforming cantonal banks can be expected to be recapitalized by the state and hence do not disappear from the sample, while weak regional banks often either go out of business or are taken over by other institutions. Secondly, between 1988 and 1990 Switzerland experienced a strong growth phase with few defaults, which probably rewarded risk-taking. Third, there is some anecdotal evidence that after having suffered huge losses during the real estate crisis of the early nineties, most Swiss banks upgraded their credit risk management and control, which until then were quite superficial and unsophisticated. This may be especially true for banks that are well monitored by their stakeholders, i.e., private banks. Finally, some cantonal banks recently went (partially) public and many cantons started to reduce the equity stake in their bank, though without withdrawing the state guarantee. This may have enhanced the influence of private shareholders on the managers of cantonal banks. Thus, one can argue that the relative importance of the moral hazard effect induced by the state guarantee increased in the last years.

In Section 3.3, I showed that in a dynamic perspective one also has to control for the role of the charter value. The trouble is that F^* is hardly observable at the single bank level, especially for those institutions – the great majority in the sample – that do not have publicly traded equity outstanding. As a proxy I employ the return on assets over the whole sample history. Accordingly, a regional bank with high franchise value is defined as a bank that, on average, has a return on assets greater than the return on assets of the median regional bank. I then compare the lending performance of cantonal banks with a sample of banks with high or low F^* respectively. The last columns in Table 3.5 show the main findings. Note that I have fewer observations in the sample of regional banks with a low franchise value. This is because underperforming banks are more likely to drop out of the sample – i.e., their time series is shorter.

Even after controlling for F^* , cantonal banks appear to be more aggressive but less profitable than regional banks. Moreover, the franchise value does not seem to significantly influence the risk attitudes of regional banks, as the coefficients of the CB dummies are almost the same in both samples. In other words, the static model of Section 3.2 seems to be a good approximation of reality. Such a conclusion clashes with the general findings of other researchers. Demsetz, Saidenberg, and Strahan (1996), for instance, observe a negative relationship between franchise value and systematic respectively bank-specific risk. How can one explain this discrepancy? On the one hand this literature uses another and probably better measure of the charter value (generally Tobin's q), which might have increased the explanatory power of the regressions. On the other, the

bank samples are quite different as they usually include only stock exchange-listed banks, which are much bigger and more sophisticated than Swiss regional banks. If the franchise value were correlated with size – e.g., due to economies of scale and business diversification benefits – and the effect of F^* on risk-taking is nonlinear, my results would no longer be inconsistent with previous evidence.

3.5.2 Procyclicality

In a second step, I examine the dynamics of the lending performance of each group. The model in Section 3.2 predicts that if private banks are really shareholder-value-maximizers subject to market discipline, their risk attitude should be more procyclical than the one of public banks that, because of bad monitoring or moral hazard, choose the H-portfolio more often. The most straightforward way to explore the procyclicality issue is through the interaction terms of the cantonal bank dummy in the translog specification. Control variables are likely to be collinear, though. Hence, interacting CB dummy_i with all the other variables may bias the coefficients and complicate their interpretation.²⁶ For that reason I estimate the translog model interacting the cantonal bank dummy only with the GDP growth rate. I call this variable GDPgrowthCB_{i,t}. In contrast, all the control variables are normally interacted. The results are shown in Table 3.6.

The figures in Table 3.6 confirm that cantonal banks assume more risks than private institutions, as β_{CB}^P is significantly positive in both samples. Moreover, in line with the model with market discipline and shareholder-value-maximizing or superficially monitored public banks, this gap becomes smaller during periods of economic expansion, as private banks are more likely to invest in the H-portfolio when π_H and π_L increase by roughly the same amount. This procyclicality effect is economically important. For instance, the model predicts that when the economy is stagnating (zero GDP growth) the difference in the ratio of provisions to total credits between a cantonal and a regional bank would be 20 basis points. During an expansion phase (3 per cent GDP growth) this difference decreases to 11 basis points. Given that the average ratio of provisions to total credits in the full sample amounts to 53 basis points, the reduction is material.

Surprising at first sight is the positive significance of GDPgrowthCB_{i,t} in (3.12), since if the investment strategy of both banks becomes more similar during good times, differences in performance should then shrink. One explanation for this result lies in the fact that the expected revenues from the H-portfolio go up more than the average revenues from the low-risk portfolio. I show in Appendix 3.7.3 that if the private bank cannot restructure its whole portfolio right away, expected revenues may increase more at the public bank when the economy is growing.

The effect on net interest revenues is in line with the model, which predicts that during economic expansion the portfolio of the public bank (unintentionally) becomes more efficient. Besides, note that for the 1987-2004 sample the CB dummy_i in (3.14) is no longer significantly positive. In

²⁶Interacting all the variables except the cantonal bank dummy produces results that are very similar to those shown in Table 3.5.

3 Is direct market discipline in banking really effective and necessary?

Table 3.6: Procyclicality

The translog-transformed values of interest revenues, provisions and net interest revenues to total loans are regressed on the cantonal bank dummy, the interaction variable GDPgrowthCB _t and a set of transformed control variables using a FGLS model with panel-specific autocorrelations. In sample F Default _{i,t} is replaced by GDPgrowth _{t-1} . The Raiffeisen group is excluded from both samples. All coefficients and standard errors are in basis points.				
Sample period:		(E) 1993-2004		(F) 1987-2004
Variable	Coefficient	(Std. Err.)	Coefficient	(Std. Err.)
<i>Interest revenues:</i>				
CB dummy _i	6.61**	(2.12)	16.60**	(2.30)
GDPgrowthCB _{i,t}	136.12 [†]	(69.64)	205.58**	(54.67)
<i>Provisions:</i>				
CB dummy _i	19.86**	(2.15)	20.78**	(2.30)
GDPgrowthCB _{i,t}	-306.19**	(99.32)	-359.48**	(70.16)
<i>Net interest revenues:</i>				
CB dummy _i	-15.79**	(3.12)	-1.57	(2.94)
GDPgrowthCB _{i,t}	439.16**	(133.00)	635.35**	(90.98)
Number of obs.		1583	2973	
Number of banks		161	241	
Significance levels : †=10% *=5% **=1%				

general, however, the findings from sample F also support the view that cantonal banks are more risky and more exposed to fluctuations in the cycle.

3.5.3 Robustness tests

Several tests were performed to assess the robustness of the results shown above. On the whole, these robustness tests corroborate the conclusions derived in the previous sections.

Sample selection

In this subsection, I explore the robustness of the results with respect to the bank sample choice. First, I exclude Banque Cantonale Vaudoise (BCV) and Banque Cantonale de Genève (BCGe) from the cantonal bank sample, as they do not enjoy a full explicit state guarantee and hence one cannot a priori exclude the possibility that market discipline does not work. Indeed, Facchinetti (2007) shows that BCV's and BCGe's bond spreads are risk-sensitive, especially when compared to the spreads of other cantonal banks. Not surprisingly, however, the lack of a full explicit state guarantee did not "prevent" both cantons from recapitalizing the respective bank after the huge losses the two institutions incurred in recent years. This is consistent with the findings of Gropp

and Vesala (2001), who provide evidence that moral hazard may be even stronger in the presence of *implicit* guarantees, as an explicit system sets a clear limit on the safety net. Moreover, I emphasized above that moral hazard is strongest at banks that enjoy a state guarantee but are controlled by private equity-holders. The exclusion of BCV and of BCGe from the sample does not materially affect the results, though. The outperformance of the regional banks becomes less pronounced, but the coefficients on the cantonal bank dummies remain significant with the same signs as in Table 3.5 and Table 3.6.

Next, I tried to provide a closer match between the regional bank and the cantonal bank sample. As noted above, the most sophisticated cantonal banks are quite different from the typical regional bank in terms of size and credit business diversification. Yet, in order to explore the effect of market discipline on lending one has to make sure that banks included in the sample do not differ in other dimensions. Because of the endogeneity problem illustrated in Section 3.4.2, I could not control for individual bank variables directly in the estimations. Instead, I excluded the smallest regional banks (total assets less than 20 million) from the sample as well as all the institutions (cantonal and regional) for which trading and wealth management activities account for more than one third of gross income. As a result, roughly 45 per cent of the observations are excluded. The findings from the Cobb-Douglas model run on this matched sample do not differ substantially from the results shown in Table 3.5. One exception is the CB dummy_i in (3.12), which is still positive but not significantly different from zero. In the translog specification the procyclicality effect becomes less pronounced, especially for the 1993-2004 period, but the main conclusions remain valid.

Explanatory variables

In a second set of robustness tests I explored different specifications with respect to the explanatory variables. First, I expanded the vector of the control variables by including lags (Libor_{t-1}) and leads (GDPgrowth_{t+1} ; Default_{t+1}).²⁷ Current interest revenues, for instance, may also depend on past interest rates, since the interest rate on loans, mortgages in particular, often remains fixed for several years. Provisions, on the other hand, tend to increase if the bank expects the creditworthiness of the borrowers to deteriorate in the future. Leads are used as proxies for expectations. Again, this expanded specification does not upset the conclusions of the base model.

Second, I analyzed the sensitivity of the results regarding alternative measures of F^* . The problem with the return on assets as a proxy for the franchise value is that it could be strongly correlated with lending performance. In other words, those regional banks that can exhibit a lending performance above average may be thought to be banks with a relatively high return on assets.²⁸ I ran the model with other proxies for F^* – net income to book capital, total dividends to total book capital – but the essential results do not differ from those shown in Table 3.5.

²⁷I do not use leads and lags in the translog model since this would considerably increase the number of coefficients to be estimated.

²⁸This hypothesis finds little support in the data, though. Among regional banks the correlation between return on assets and interest revenues, provisions and net interest revenues amounts to -0.08, -0.42 and 0.08 respectively.

3 Is direct market discipline in banking really effective and necessary?

Additional measures of risk and performance

In a final set of robustness tests, I ran the translog model for additional, more traditional measures of risk and performance. The results for the full sample and the matched sample are shown in Table 3.7.

Table 3.7: Additional measures of risk and performance

The translog-transformed values of the variables listed in the first column are regressed on the cantonal bank dummy and a set of transformed control variables using a FGLS model with panel-specific autocorrelations. In the matched sample the smallest regional banks as well as banks with significant trading and wealth management activities are excluded.		
Dependent variable	Coefficient on CB dummy _i	
	Full sample	Matched sample
Return on assets	Negative**	Negative
Return on equity	Negative	Positive*
Sharpe ratio (RoA)	Negative**	Negative**
Sharpe ratio (RoE)	Negative**	Negative**
Share of first rank mortgages [‡]	Negative**	Negative
Leverage	Positive**	Positive**
[‡] First rank mortgages are mortgages with a low loan-to-value (LTV) ratio.		
Significance levels : †=10% *=5% **=1%		

The findings are in line with the conclusions derived from the analysis of lending performance. Depending on the sample and on the figure one looks at (return on assets or return on equity), cantonal banks may be more or less profitable in absolute terms. However, after controlling for risk (Sharpe ratio), regional banks significantly outperform public banks. Finally, regressions on other measures of risk (quality of mortgages, leverage) suggest that cantonal banks indeed tend to assume more risk. As noted in the introduction to Section 3.4, however, some caution is needed with respect to those results.

3.5.4 Discussion

The basic findings above suggest that while cantonal banks have to set aside more provisions each year, their total net interest revenues (non risk-adjusted) tend to be lower than net interest revenues of other bank groups with an a priori similar business model, especially in the last decade. In other words, public banks assume more risks than private banks, but they make less money out of it, even in absolute (i.e., non risk-adjusted) terms. These findings are in line with the evidence provided by Iannotta, Nocera, and Sironi (2007) for large European banks. According to the static model (see Table 3.2), this combination of results implies that the market can directly discipline private banks and, as a corollary, that the state guarantee leads to an inefficient allocation of resources. That cantonal banks appear to be more exposed to fluctuations in the cycle corroborates this conclusion.

Supporters of the state guarantee and of cantonal banks would counter the statement above with the following argument: “It is true that public banks assume more risks. But they invest in projects with a high economic value. That net interest revenues are low when compared with other bank groups can be explained by the fact that cantonal banks offer more attractive contracts – i.e., lower risk-adjusted interest rates – because they do not maximize shareholder value. Private banks are too conservative and risk-averse and do not finance projects or firms that have a relatively high probability of default but, when successful, make an above average contribution to economic value added”. That some bank groups tend to be too restrictive in their credit policy is a widely held opinion nowadays, in particular in the context of SMEs. In Section 3.3 we saw that private banks may indeed be too risk-averse when their franchise value is high enough, even in the absence of market discipline.

However, several cases can be made out for refuting or at least weakening this line of argumentation. First, the franchise value does not appear to significantly influence the credit policy of Swiss regional banks. Thus, the static framework, in spite of its simplicity, seems to be a good first approximation of reality. Secondly, empirical studies on the Swiss banking sector – e.g., Rime and Stiroh (2003) or Bichsel (2006) – provide evidence that cantonal banks are not less (or more) efficient than other bank groups and do not have a significantly different credit pricing policy. Moreover, do public commercial banks really refrain from maximizing shareholder value (or at least profits) and focus on supporting the local economy? Further research is needed to answer this question, but press releases issued by cantonal banks on their annual results do not give the impression of management being indifferent to profitability.²⁹ Finally, one can argue that if public institutions actually offer more favorable contract terms than other bank groups, i.e., the borrower retains a bigger part of the rent, everybody who needs funds would try to go to a cantonal bank. Consequently, the presence of public banks can be seen as fuelling competition and their market share should grow continuously – and with no financing problems because the state guarantees gives cantonal banks easy access to cheap money. However, this does not seem to correspond to reality. Bichsel concludes that “they neither adopt a customer friendly conduct nor intensify competition in any special way”. Furthermore, in the last ten years cantonal banks have constantly made up one-third of the domestic credit market and regional banks 10 per cent (see Figure 3.1). Only the Raiffeisen group has been able to strengthen its position, without being excessively aggressive, however (see Table 3.5). That between 1987 and 1993 the market share of cantonal banks increased while that of regional banks declined can be attributed to the fact, mentioned earlier on, that in these years Switzerland experienced a real estate crisis which forced domestic banks to write off credits worth several billion Swiss francs. A lot of regional banks ran into difficulties and were taken over by bigger institutions. For instance, between 1992 and 1993 nineteen regional banks disappeared from the Swiss banking landscape; five of them were acquired by cantonal banks.

²⁹La Porta, Lopez-De-Silanes, and Shleifer (2002) find a negative relationship between government ownership of banks and economic growth in a cross-section of countries. This can be interpreted as further evidence that public banks do not improve economic performance.

3 Is direct market discipline in banking really effective and necessary?

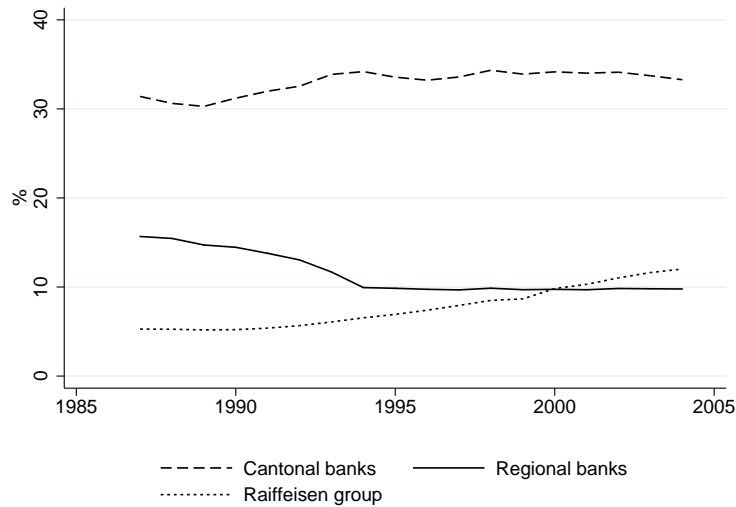


Figure 3.1: Market share in the Swiss credit market

What if cantonal banks are benevolent only towards those borrowers that other banks do not finance, but otherwise behave like private banks and maximize profits? Since the study by Bichsel focuses on the mortgage business, his conclusion may indeed not be representative for other credit segments, SME financing for instance. Yet, it is unclear why cantonal banks should “subsidize” some borrowers and not others, given also that the maximization of economic value added – the typical public mandate – is not necessarily incompatible with profit-maximizing.³⁰ Moreover, we then have to ask ourselves *why* other banks do not lend money to those borrowers. Given that the franchise value seems not to play a role for regional banks, there are two possible answers: Either (a) those borrowers do not have an adequate risk-return profile, i.e., they are bad risks; or (b) regional and Raiffeisen banks are not able to lend money to those borrowers (e.g., because of their modest size) while at the same time big banks have a high franchise value that makes them too risk-averse. Hypothesis (b) seems to be rather weak for the following reasons. First, it can hardly hold for SMEs, the credit segment for which it has often been argued that private banks are excessively risk-averse. Second, there is no reliable evidence that UBS and CSG avoid risk-taking in the domestic loan market: Although their market share decreased somewhat in recent years, big banks are the dominant players in many domestic loan segments.

If (a) is true, one question remains unanswered: Why should cantonal banks actually assume those bad risks? In Section 3.2, I identified two possible rationales: a moral hazard effect induced by the state guarantee, and relatively superficial monitoring by the state. Unfortunately the two effects are observationally equivalent in my econometric model. I cannot, therefore, give a definitive answer to this question, which is thus left for future research.

³⁰Remember that this is generally not the same as shareholder-value-maximizing. See Section 3.2.

3.6 Conclusions

There is sufficient empirical evidence that investors require risk-adjusted compensation for the money they lend to banks, for instance in the form of deposits or bonds. The main aim of this study was to find out whether the market is able to directly discipline banks through this risk-sensitive supply of funds. The underlying idea was to compare the lending performance of banks with a state guarantee, whose financing costs are expected to be risk-insensitive, with those of other bank groups. I show with a simple model that because of moral hazard induced by the state guarantee and/or less strict monitoring by the equity-holder (the state), public banks may pursue more risky though less efficient strategies than banks that are privately owned and subject to market discipline. On the other hand, the state guarantee may prevent high-charter-value banks from being excessively risk-averse and thus possibly generate economic value added, which would justify the existence of public banks in developed countries. The first hypothesis finds support in Swiss data. The paper provides empirical evidence that cantonal banks assume more bad risks – measured by the amount of new provisions built up each year – than other bank groups active in similar market segments. However, their net interest income before financing costs is lower, indicating that the risks they assume are not adequately rewarded. This seems to be true especially in recent years, as Swiss banks started to improve their credit risk management framework after the real estate crisis of the early nineties and cantons started to reduce their stakes in their own banks without reducing the scope of the state guarantee. In addition, the argument that this result is due to the fact that cantonal banks have objectives other than maximizing shareholder value appears to be rather weak. Hence, one can conclude that the market exercises some direct discipline on private banks. As a corollary, public banks appear to operate inefficiently either because of moral hazard or less strict monitoring (or both). Perhaps it is just an unfortunate coincidence, but in the last decades the taxpayers of several Swiss cantons have had to recapitalize their state bank after it incurred huge losses due to its credit business.

3.7 Appendix

3.7.1 Complete results of the Cobb-Douglas model

The complete estimation results for sample A (cantonal and regional banks between 1993 and 2004) are shown in Table 3.8.

3 Is direct market discipline in banking really effective and necessary?

Table 3.8: Complete results of the Cobb-Douglas model

Cross-sectional time series FGLS regressions with panel-specific autocorrelations. Coefficients and standard errors are in percent.				
Dependent variable: Explanatory variable	Gross interest income _{i,t} Coefficient	Provisions _{i,t} Coefficient	Net interest income _{i,t} Coefficient	
GDP _{grow} _t	-15.72** (0.58)	-0.90* (0.35)	-15.59** (0.60)	
Libor _t	60.92** (1.38)	6.50** (0.88)	55.62** (1.41)	
YieldCurve10Y _t	28.27** (1.80)	2.19* (1.08)	24.78** (1.76)	
Default _t	1.37 (1.31)	0.78 (0.86)	2.33† (1.30)	
REcrisi _{i,t}	0.03 (0.03)	0.11** (0.01)	-0.09** (0.02)	
Area _{i,t}	-2.56** (0.53)	-0.46 (0.34)	-1.74** (0.49)	
Pop _{i,t}	17.18** (3.57)	21.01** (2.78)	-15.28** (3.00)	
BankNr _{i,t}	321.33** (12.50)	27.99** (6.88)	236.27** (9.84)	
CB dummy _i	0.06* (0.03)	0.20** (0.01)	-0.11** (0.03)	
DataBreak dummy _t	0.04† (0.02)	-0.03* (0.02)	0.04 (0.03)	
Constant	2.43** (0.12)	0.01 (0.08)	2.43** (0.11)	
Number of obs.	1583	1583	1583	
Number of banks	161	161	161	
Log-likelihood	7273.3	7654.0	7043.6	
χ^2	14724.5	1819.0	26208.9	

Significance levels : †=10% * =5% **=1%

3.7.2 Risk-taking incentives and debt

The vega of a derivative, in our case bank's shareholder value, measures the sensitivity of the derivative value with respect to the volatility of the underlying, in our case bank asset risk. According to the Black and Scholes (1973) model for a European call option E on a non-dividend-paying asset A with a strike price equal to D (in our case the amount of debt), vega V is approximately³¹ given by

$$\begin{aligned}\frac{\partial E}{\partial \sigma} &= V = A\phi(d_1) \\ &= A\phi\left(\frac{\ln A - \ln D + r_f + 0.5\sigma^2}{\sigma}\right),\end{aligned}$$

where $\phi(\cdot)$ is the density of the normal distribution and σ the volatility (risk) of the asset. Time to maturity is normalized to 1 for simplicity. Notice that

$$\frac{\partial V}{\partial D} = -\frac{A\phi'(d_1)}{\sigma D} > 0$$

when $\phi'(d_1) < 0$, that is when

$$d_1 = \ln A - \ln D + r_f + 0.5\sigma^2 > 0,$$

which is always satisfied if $A > D$, i.e., when the bank is solvent. This means the marginal utility of assuming more risk, i.e., the marginal increase in shareholder value, goes up with the amount of debt (including interests) to be paid at the end of the period.

3.7.3 Effect of the cycle on expected revenues

Suppose that the private bank is subject to market discipline while the public bank either maximizes shareholder value, which generates moral hazard, or is subject to less strict monitoring by its equity-holder. In both cases the public institution chooses the H-portfolio more often than the private bank, as shown in Section 3.2.1. If the general conditions change, it may be the case that both banks would like to switch immediately from one portfolio to the other. In reality, however, banks – in particular the smallest ones, which are not active in the credit derivative and securitization market – cannot restructure their portfolio within a short period of time. Assume for instance that the public (private) bank is currently investing a fraction λ' (λ'') of its assets in the H-portfolio and the rest in the low-risk portfolio, with $\lambda' > \lambda''$. Then, if π_H and π_L go up by the same amount $d\pi$ between t and $t + 1$, we see that the expected revenues of the public bank increase by

$$\lambda' R_H d\pi + (1 - \lambda') R_L d\pi + (\pi_H R_H - \pi_L R_L) d\lambda'. \quad (3.15)$$

³¹Since the Black-Scholes model assumes constant volatility, calculating vega in their framework represents an approximation. See Hull and White (1987).

3 Is direct market discipline in banking really effective and necessary?

For the private bank, the growth in expected revenues amounts to

$$\lambda'' R_H d\pi + (1 - \lambda'') R_L d\pi + (\pi_H R_H - \pi_L R_L) d\lambda'', \quad (3.16)$$

where $d\lambda'$ and $d\lambda''$ represent the respective change in the portfolio composition between t and $t + 1$. Since $\lambda' > \lambda''$ and $R_H > R_L$ (3.15) is always higher than (3.16) for small values of $d\lambda'$ and $d\lambda''$. Consequently, if banks cannot instantaneously restructure the whole investment portfolio, the public bank will experience a higher growth in revenues during periods of economic expansion. Note that in the moral hazard case λ' always equals 1.

4 The endogeneity pitfall in using market signals as triggers for prompt corrective action

“[M]arkets ... are extremely efficient, effective and timely aggregators of dispersed and even hidden information.” (Defense Advanced Research Projects Agency, US Ministry of Defense, July 2003)

How¹ should a bank supervisory authority react to seeing the yields on an individual banking institution’s debt soar? To begin with, the supervisory authority would probably be concerned. After all, the yield of a debt instrument – as far as it exceeds the yield of a risk-free asset – measures the market perception of the issuer’s default risk. A strong increase in the yield would signal that the market has bad news about the solvency of the debtor. Such news may not be otherwise available to the supervisor. Researchers have found that market prices and quantities can in fact contain useful information not yet known by supervisors (see below).

Disregarding market signals therefore can hardly be an optimal strategy for a supervisor. This is recognized by an increasing reliance by supervisors on market data – on indirect market discipline, in the terminology of the “Basel II” agreement among bank supervisors and central banks.² The former chairman of the Federal Reserve Board Greenspan has admitted the potential value of market information:

“Significant changes in a banking organization’s debt spreads, in absolute terms or compared with peer banks, can prompt more intensive monitoring of the institution.” (Greenspan, 2001)³

As that statement also illustrates, supervisory authorities have been reluctant to pin down their reactions to market movements in advance. Proposals made by Calomiris (1999) or Evanoff and Wall (2001) suggesting a direct link between a bank’s risk premium over the market rate of interest and supervisory intervention have met with skepticism on the side of bank supervisors. “I do not need a spring gun”, a Swiss supervisor remarked to one of the authors of this paper.

Central bankers have known the issue of price-based intervention for long. It arises, for example in the context of exchange rate bands. Over the last few years it has been discussed under the

¹Urs W. Birchler is the co-author of this paper.

²*Direct* market discipline refers to a risk-sensitive supply of funds by investors. *Indirect* market discipline means that supervisors benefit from the information content of market prices in their assessment of bank risk.

³For further references see Bond, Goldstein, and Prescott (2006).

4 The endogeneity pitfall in using market signals as triggers for prompt corrective action

heading: “Should central banks burst bubbles?” Both, advocates, like Roubini (2006), and skeptics, like Posen (2006) seem to agree that asset prices may contain information useful for monetary policy.

More recently, interest in prices as guides for action came from an unlikely corner. In 2003, an agency of the US Ministry of Defense disclosed plans for “political analysis market”, a futures exchange where traders could speculate on events like terrorist attacks or political assassination (for details, see Wolfers and Zitzewitz, 2004). This market would have been introduced with the explicit goal of learning from prices. The proponents of the idea argued that markets aggregate dispersed information (see our introductory quotation) and could help to predict and, hopefully, to prevent terrorist attacks.

After becoming public the project was quickly abandoned amid a storm of political and moral opposition.⁴ There was not much of an economic debate about whether the plan might have worked at all. Not even those few economists⁵ who voiced their views asked whether prices could at the same time *predict* an event and help to *prevent* it.

The present paper tries to close this gap. We use a dynamic rational expectations model to investigate the trading strategy of investors endowed with private information about the probability of failure of a particular bank. In this setting a supervisory authority tries to extract information from the observed market prices. It uses this information to optimize costly action to prevent bank failure. We argue that this use of market signals for policy purposes may fundamentally change their information content or may even lead to a market breakdown. It is easy to see why. Prices, in the ideal case, reflect all information available to the market, including beliefs about any future policy that may affect economic fundamentals. Technically speaking, prices are endogenous to policy. If, in turn, policy reacts to prices, both prices and policies mutually depend on each other. There is a “double endogeneity problem”.

In the context of bank supervision this problem has largely been overlooked both by advocates and opponents of price-based action. This seems odd, as it has long been recognized in the context of exchange rate target zones, for example by Krugman (1991) who also shows that an equilibrium may exist despite the mutual endogeneity of prices and actions.

In Section 4.1 we review the literature on the information content of market prices for bank supervisors. In Section 4.2 we introduce the double endogeneity effect informally by way of examples. In Section 4.3 we develop a formal model of asset pricing under price-based intervention. In Section 4.4 we discuss the results and in Section 4.5 we draw the policy conclusions.

⁴Ironically, a betting market correctly predicted the imminent resignation of the person responsible for the project.

⁵An example is Hal Varian, “A Good Idea With a Bad Press”, *The New York Times*, 31 July 2003.

4.1 The informational content of prices for bank debt and equity

It has long been recognized that prices aggregate “the dispersed bits of incomplete and frequently contradictory knowledge which all the separate individuals possess”, as von Hayek (1945, p. 519, 525) put it. Grossman (1976) later formalized the idea that market prices are a sufficient statistic for underlying information, at least under simplifying assumptions. Complications arise when markets are incomplete (Brunnermeier, 2001), in the presence of *noise* (Hellwig, 1980) or when information is *endogenous* (Grossman and Stiglitz, 1980; Verrecchia, 1982). An observed price structure thus may not be a perfect mirror of the underlying fundamental conditions, but still a useful source of information.

Markets add the fire of self interest to the fuel of statistics: “market participants have an incentive to look through reported accounting figures to the real financial condition of a bank and to price a bank’s securities based on their best estimates of the distribution of the security’s future cash-flows” (Flannery, 2001). Although supervisors can collect confidential data, market information may be just as reliable.

Several empirical studies have examined the information content of market data on bank risk for the US (Flannery, 1998; Berger, Davies, and Flannery, 2000; DeYoung, Flannery, Lang, and Sorescu, 2001; Evanoff and Wall, 2002; Swidler and Wilcox, 2002; Fan, Haubrich, Ritchken, and Thomson, 2003; Krainer and Lopez, 2004; Krishnan, Ritchken, and Thomson, 2006; Curry, Elmer, and Fissel, 2007), for European countries (Persson, 2002; Sironi, 2003; Cannata and Quagliariello, 2005; Gropp, Vesala, and Vulpes, 2006), for Japan (Brewer, Genay, Hunter, and Kaufman, 2003), and for emerging countries (Bongini, Laeven, and Majnoni, 2002; Sy, 2002).

The different studies show that in industrial countries markets, particularly those for equity and for subordinated debt (Basel Committee on Banking Supervision, 2003), are sufficiently developed to provide meaningful information for supervisors. Several researchers argue that market data contain information which is *not yet* part of confidential supervisors’ information. Conversely, market data do not reflect *all* information available to supervisors. In other words, the empirical evidence suggests that supervisory information and market information are *complementary* sources of bank supervisory intelligence.

One might conclude that supervisors should not neglect the information contained in market prices. However, the evidence cited stems from a period in which supervisors did probably not react much to changes in market data. It may not carry forward to an environment where markets expect supervisors to *react* to market data.

4.2 The double endogeneity problem

4.2.1 A potential paradox

Let us start from an example. Assume that banking supervisory authority, following a suggestion by Calomiris (1999), declares to intervene whenever the risk premium a bank pays on subordinated debt (the difference in yield to a comparable treasury bond) hits some threshold. Let this threshold be a 5 per cent p.a.. Assume further that an intervention (a mandatory recapitalization, say) is always successful in the sense that it prevents the bank from defaulting.

What would be the consequence? Rational investors would anticipate that whenever a bank's subordinated yield spread hits the 5 per cent ceiling, recapitalization would follow. After recapitalization, the bank would be quite safe again. No investor would thus sell subordinated debt at a discount corresponding to a yield spread of 5 per cent. Thus, the spread would never hit the threshold, and the authority would never intervene – but then the bank may indeed fail, and the spread should reflect that failure.

This paradox is explained by the informational efficiency of the market. In an informationally efficient market all available information is reflected in the price. This includes information or beliefs about supervisory reactions to observed prices. The price – here: the implied risk premium – therefore does not reflect the bank's risk pre-intervention, but post-intervention. Prices, in other words, are *endogenous* to expected policy reactions. An authority who misreads market prices as signals of pre-intervention probabilities would therefore be misguided.

The example is reminiscent of the well-known case of an exchange rate target. In the context of exchange rate intervention, the endogeneity effect has long been known. Krugman (1991) for example pointed out that exchange rate target zones had an impact on exchange rates long before these hit any of the intervention points. In fact, if the target band is perfectly credible, the central bank never has to intervene.

4.2.2 Fragile equilibrium

The exchange rate example also illustrates that even in the face of endogenous action, there may be an equilibrium. Market participants try to anticipate the authority's action, while the authority tries to anticipate market participants' reasoning. In equilibrium, both parties are right and hold consistent beliefs. In the context of bank supervision the logic behind an equilibrium can be shown with the following arbitrary example.

The supervisory authority again observes the risk premium banks pay on subordinated debt. It is common knowledge that the authority acts according to the following scheme: (i) from the observed risk premium the authority infers the probability with which the bank would fail in the absence of supervisory intervention; (ii) the authority intervenes in a way to reduce the perceived probability of failure (without intervention) by half.

The authority may observe a risk premium for a bank of 3 per cent. With risk-neutral investors

and a zero recovery rate this would roughly correspond to a 3 per cent probability of failure. This would suggest that the market believes the bank has a 6 per cent probability of failure. The authority would take corrective action to reduce this probability by half, leaving the bank with a 3 per cent post-intervention chance of failure. This is exactly what market participants expected when they traded the bank's debt. No party regrets their decision in the light of other parties' decisions; the market thus is in *equilibrium*.

The equilibrium price fully reveals the bank's failure probabilities, both with and without supervisory intervention. However, while it directly reflects the post-intervention probability (of 3 per cent), it only indirectly reveals the pre-intervention probability (of 6 per cent).

Note that in the example the authority did not reduce the bank's risk of failure to zero. A "zero tolerance" policy would kill the message (the information content of the market price). So would any policy to achieve a fixed target probability of failure. The post-intervention probability of failure must be an increasing function of its pre-intervention probability if an equilibrium should exist.

4.2.3 Previous literature

We have already mentioned that the mutual endogeneity of prices and price-based action was recognized in the context of exchange rates by Krugman (1991). In the debate about whether central banks should react to asset price developments (to bubbles, in particular), Cecchetti, Genberg, and Wadhvani (2002, p. 5) for example note that the mere expectation of a central bank's intent to react to asset prices would already change their dynamics: "if it were known that monetary policy would act to 'lean against the wind' in this way, it might reduce the probability of bubbles arising at all".

In the context of bank supervision the double endogeneity effect was long overlooked. DeYoung, Flannery, Lang, and Sorescu (2001) were led to the effect by some puzzling empirical evidence: They found that after an unexpectedly poor outcome of an on-site examination bond spreads of financially troubled banks *decreased*, rather than increased. They conclude that a poor result of a supervisory examination was more likely to trigger prompt corrective action and that "the anticipated regulatory response frequently dominates the information's implications about current bank conditions." (DeYoung, Flannery, Lang, and Sorescu, 2001, p. 902). Shin (2002) mentions that banks' asset values were affected by the endogenous response of authorities during a currency crisis.

An explicit discussion of the endogeneity effect in the context of bank supervision can be found in Birchler and Facchinetti (2007). Bond, Goldstein, and Prescott (2006) offer a more formal treatment of some aspects. In their model the action set of the authority is binary, i.e., it can either intervene or not intervene. They show that under reasonable parameter values, there may be multiple equilibria or no equilibria at all. A key determinant for the existence or efficiency of equilibria is the difference between the authority's and the market's prior uncertainty about bank

fundamentals (the “information gap”).

Another model of price-based intervention was developed by Lehar, Seppi, and Strobl (2006). Using a binary intervention space as well they find that the supervisory use of market prices may distort information and incentives. The focus of the paper is on the interaction between policy response and *information acquisition*, rather than the efficient representation of *existing* information in prices.

An important difference between this literature and our paper is the modeling of supervisory action. While both Bond, Goldstein, and Prescott (2006) and Lehar, Seppi, and Strobl (2006) only leave a binary “yes-or-no” action to the authority, in our model the supervisory authority’s action space is *continuous*. This is not only more realistic. More importantly, in a continuous action setting, a unique equilibrium is much more likely to exist than under binary action. Under a continuous intervention space with monotonous intervention cost the unique equilibrium is also *information revealing*.

The special case of binary action in our model is captured by use of a “quasi-fixed cost” of intervention. In the presence of a quasi-fixed cost the optimal action level “jumps” at some critical market price. As a result, we find effects similar to those in the binary setting of Bond, Goldstein, and Prescott (2006). Their point of departure, “the fact that prices *affect* and *reflect* the supervisor’s action at the same time makes the analysis of equilibrium outcomes [...] quite hard” (Bond, Goldstein, and Prescott, 2006, p. 4), is also a good “leitmotiv” for our model.

4.3 The model

4.3.1 Assumptions

Outcomes

Our simple economy has two final outcomes. These are denoted by $\omega \in \{1, 0\}$ indicating whether a certain negative “event” has been avoided or not. We will think of the event as the failure of a bank.

The exogenous prior probability of the negative event “bank failure” is $\Pr(\omega = 0) = \pi_0$. With the complementary probability $\Pr(\omega = 1) = 1 - \pi_0$ the bank remains solvent. We assume that nature draws π_0 from a uniform distribution

$$\pi_0 \sim U \left[\frac{1}{2} - \sigma; \frac{1}{2} + \sigma \right],$$

with $\sigma \leq \frac{1}{2}$.

The authority

An authority prefers the bank to remain solvent. The risk-neutral authority considers a bank failure as a deadweight social loss of $v(\omega)$. We normalize the loss in case of solvency to zero, i.e.,

$$v(0) = v \geq v(1) = 0.$$

Based on its beliefs (to be modeled below) about the true value of π_0 the authority can take some action $a \in [0, \infty]$ to influence the posterior probability of the event π . In the context of bank supervision the authority may, for example, increase capital requirements or change bank management in order to reduce the probability of a bank default. The posterior likelihood that the bank fails is therefore a function of the prior probability of default π_0 and of the intensity of intervention, that is

$$\pi = \pi(\pi_0, a).$$

For simplicity we assume that:

$$\pi = \frac{\pi_0}{1+a}. \quad (4.1)$$

The cost of intervention is given by

$$c = c(a),$$

with $\frac{\partial c}{\partial a} > 0$, $\frac{\partial^2 c}{\partial a^2} \geq 0$ and $c(0) = 0$. In addition $c(a)$ has a quasi-fixed costs component, that arises when $a > 0$. We assume that:

$$c(a) = ka + K(a), \quad (4.2)$$

where

$$K(a) = \begin{cases} K & \text{if } a > 0 \\ 0 & \text{otherwise.} \end{cases}$$

This implies that the authority takes action (i.e., chooses $a > 0$) only when the benefits of intervention are high enough so to justify the occurrence of the quasi-fixed costs. In Section 4.3.4 we show that this is the case when the prior probability of bank failure exceeds some threshold. As a tie-breaker we assume that in the border case where the authority is indifferent it does not intervene.

The quasi-fixed cost implies a “jump” in the cost function. For simplicity we assumed that this jump occurs at $a = 0$. One could also assume that a jump occurs at a positive level of a , for example at the point when the authority’s intervention reaches a dimension that makes it public, assuming that the authority prefers “covert” interventions.

In the context of banking supervision, the quasi-fixed costs setting appears particularly realistic. In fact, it is reasonable to assume that supervisors dislike interfering in banks’ business. This is particularly true for measures that are publicly observable, like a mandatory issue of capital, a dividend reduction or a management removal as such measures may have additional negative implications, like public criticism or even panic reactions by depositors. Proposals to tie supervisory

intervention to market prices (Calomiris, 1999; Evanoff and Wall, 2001) exactly aim at forcing authorities to overcome quasi-fixed cost barriers.

Individuals

The authority cannot observe π_0 directly. There is, however, a mass of homogeneous, risk-neutral individuals. Some of these (the “insiders”) can observe π_0 . While the realization of π_0 is private knowledge of insiders, everything else, including the distribution of π_0 , is common knowledge. For reasons to be explained below, we call the other, uninformed individuals “noise traders”.

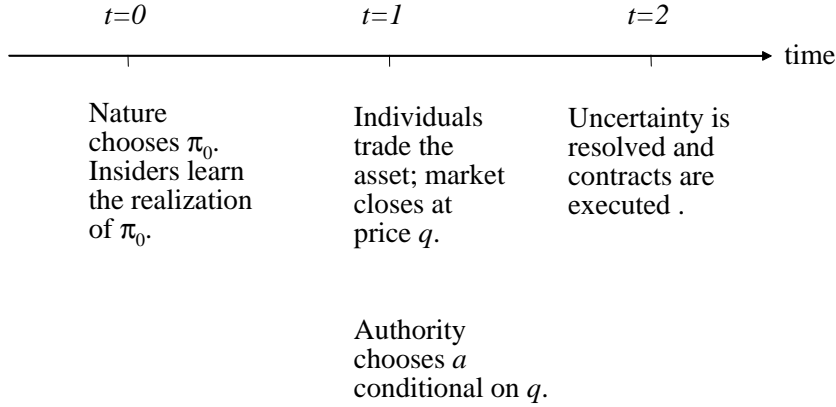
Each individual, insider or uninformed, is endowed with some amount of the numeraire.

The market

There is an asset which pays one dollar if the negative event does not happen ($\omega = 1$), that is if the bank remains solvent, and nothing if the event happens ($\omega = 0$), that is if the bank fails. The asset is traded in a market free of operational cost. The asset can be bought as well as sold short. However, if investors sell the asset short, they must pledge numeraire as collateral. Block trades are not allowed. Nature randomly selects the order investors can trade. When it is their turn, investors submit the desired order. The price at which the asset trades is denoted by $q \in [0; 1]$. q is publicly observable.

For most of the paper it would be sufficient to think of a frictionless market in which informed individuals immediately drive the price to its equilibrium value. Yet, we prefer to use a dynamic model of market microstructure with sequential trades. Such a setting is more close to reality and the spirit of Calomiris (1999), where authorities look at the evolution of asset prices over time. Secondly, we believe that a model of sequential trades makes the derivation of the main results more intuitive.

In addition, we explicitly model the supply side of the market. As Milgrom and Stokey (1982) have shown, the existence of informed individuals, if commonly known, prevents trading: Uninformed individuals would never rationally speculate against informed individuals. Kyle (1985) argues that a positive amount of trading can occur if there are a sufficient number of noise traders who trade for reasons other than speculation. In order to allow for a positive amount of trading even in the absence uninformed investors, we assume that the authority acts as a market maker offering to buy and sell the asset at the same price. The authority stands for the community of noise traders in the sense of Kyle (1985). Market making often imposes a loss to the authority. We assume that the authority considers such loss as a sunk cost required to get valuable information from observing the equilibrium price. From a social point of view, the authority’s loss is compensated by individuals’ profits. Anyway it could be recovered by a lump sum tax. Our assumptions imply that there is in principle no restriction on the price setting process, as long as it optimizes the information extraction. As a corollary, we can assume without loss of generality that the bid-ask spread is always zero.

**Figure 4.1:** Time line

The market starts with the authority posting a price reflecting the prior probability π_0 . At that price the authority gets either a buy order ($x > 0$) or a sell order ($x < 0$). After observing the direction of the trade, the authority adjusts its expectations about π_0 – down after a buy order, up after a sell order – and the price of the security. This process goes on until $x = 0$ or until the market closes.

In a general context, the current price of the security as well as its volume traded represent the market signals, \mathbf{M} , available to the authority. Since the authority acts as market maker and sets the price on the basis of the order book, we have $\mathbf{M} = \{x_1; x_2; \dots\} = X$. The price becomes a sufficient statistic of the information set of the authority. At the end of the “day”, when all individuals have posted their orders, the market closes with price q . As q exhausts the information set of the authority, we can use it instead of \mathbf{M} .

Time line

Events in the model are summarized by the time line represented in Figure 4.1. In $t = 0$, insiders learn the true realization of π_0 . In $t = 1$, the security is traded in the market. Still in $t = 1$, the authority, upon observing the market signals, can take its action a to reduce the probability of default. In $t = 2$, finally, nature decides whether the bank defaults or not. Individual investors get their respective payoffs from having bought or sold the asset.

4.3.2 Optimal strategies

The authority and individuals (among which the insiders) play a sequential game. Individuals move first (by trading), followed by the authority (choosing an action level). Solving the game backwards we first look at the authority’s maximization problem.

The authority

If the authority knew the prior probability of the event, π_0 , it would minimize expected loss by solving

$$\min_a L = \pi(\pi_0, a) v + c(a). \quad (4.3)$$

As the authority cannot observe π_0 , it has to rely on indirect evidence. The authority tries to infer the private knowledge of insiders and hence the true value of π_0 from observed market signals. The resulting beliefs are denoted by π_0^R . Throughout the paper we denote beliefs by superscripts (R for the authority, I for individuals). π_0^R is a commonly known function of market signals, \mathbf{M} , and a is a function of π_0^R , i.e., $\pi_0^R = \pi_0^R(\mathbf{M})$ and $a = a(\pi_0^R)$. As mentioned above, the authority revises its beliefs π_0^R downward after observing a buy order and upwards after a sell order. Hence, $\frac{\partial \pi_0^R}{\partial x} < 0$.⁶

Relying on such beliefs the authority solves

$$\min_a L = \pi^R(\pi_0^R(q), a) v + c(a), \quad (4.4)$$

where π^R is the rationally believed posterior event probability, that is the expected default probability resulting from taking action a given the beliefs π_0^R . Remind that q sums up the authority's information \mathbf{M} .

Under the specification chosen in (4.1) and (4.2) the authority's problem reads:

$$\min_a L = \frac{\pi_0^R(q)}{1+a} v + ka + K(a),$$

where $K(a)$ is the quasi-fixed component of the cost function, with $K(0) = 0$ and $K(> 0) = K$. As the authority takes action ($a > 0$) only if the expected gains of intervention exceed the quasi-fixed costs, i.e., if

$$K < \hat{K} = \pi_0^R v - \frac{\pi_0^R v}{1+a} - ka = a \left(\frac{\pi_0^R v}{1+a} - k \right), \quad (4.5)$$

the first-order condition (FOC) is

$$\begin{aligned} a &= \left(\frac{\pi_0^R v}{k} \right)^{\frac{1}{2}} - 1 \text{ if (4.5) is satisfied;} \\ a &= 0 \text{ otherwise.} \end{aligned} \quad (4.6)$$

Competitive individuals

Individuals maximize their expected profits from trading the asset. Since the asset pays one dollar if the bank ends up solvent and nothing otherwise, competitive risk-neutral investors buy (sell) the

⁶In other words, informed investors always trade in the "right direction" and do not try to fake the authority. We will see that this assumption is compatible with our equilibria.

asset as long as its price q is below (above) $1 - \pi^I$, individuals' believed posterior probability of solvency. Market equilibrium is characterized by:

$$q^C = 1 - \pi^I.$$

Building their beliefs, rational investors take two kinds of information into account. The first is the observed prior probability π_0 . The second is the anticipation that the authority will react on the market price it will observe. This is relevant for individuals, as the authority's choice of a has a direct effect on the value of the asset. Thus

$$\pi^I = \pi^I(\pi_0, a^I(q)), \quad (4.7)$$

where $a^I(q)$ denotes individuals' belief regarding the action the authority will take upon seeing the market signals. A rational belief regarding $a^I(q)$ is:

$$a^I(q) = \arg \min L,$$

that is, individuals anticipate that the bank minimizes expected loss, as expressed in (4.4).

Monopolistic individuals

Things are a little different when there is only one insider (or a few insiders) endowed with a substantial amount of the numeraire. Since block orders are not allowed, this monopolist usually does not trade only once. When the insider accumulates inventories i , however, a trade-off emerges. To see this, suppose for instance that the informed investor has already built up some long positions ($i > 0$), but q is still lower than $1 - \pi^I$, that is the believed fundamental value of the security is higher than its price. The insider is tempted to buy some more contracts, but by doing so she would signal that π_0 is actually lower than π_0^R . The reaction of the supervisor would be to intervene more moderately, which would increase the posterior probability of the default and consequently reduce the fundamental value of the contracts the insider already bought. In other words, at some price $q < 1 - \frac{\pi_0}{1+a^I}$ the insider stops trading.

More specifically, this is the case when the marginal change in the value of the portfolio of the insider $i(1 - \pi^I)$ equals the costs of the next trade dx , with dx small. That is, when

$$(1 - \pi^I) dx - i \frac{\partial \pi^I}{\partial x} dx = q^S dx$$

or

$$q^S = 1 - \pi^I - i \frac{\partial \pi^I}{\partial x}.$$

4.3.3 Equilibrium: Benchmark cases

Perfect information

Under perfect information, the authority can directly observe π_0 . It then chooses a according to (4.3). The problem of the authority reads

$$\min_a L = \frac{\pi_0 v}{1+a} + ka + K(a)$$

and the FOC is

$$\begin{aligned} a &= \left(\frac{\pi_0 v}{k} \right)^{\frac{1}{2}} - 1 \text{ if (4.5) is satisfied;} \\ a &= 0 \text{ otherwise.} \end{aligned}$$

The level of action that solves the problem of the authority under perfect information is first best and we call it a^* .

No information

Under no information neither the authority nor individuals can observe the prior event probability π_0 . What action would the authority choose?

The rational belief about π_0 would be its expectation

$$\pi_0^R = E(\pi_0) = 0.5.$$

Under this belief, the optimal action is:

$$a = \left(\frac{v}{2k} \right)^{\frac{1}{2}} - 1$$

if (4.5) is satisfied and 0 otherwise. Therefore the equilibrium price of the security with no informed individuals, is

$$q = 1 - \left(\frac{k}{2v} \right)^{\frac{1}{2}}$$

if the authority intervenes, i.e., if K is low enough, and 0.5 if it does not.

4.3.4 Equilibrium: Asymmetric information and competitive insiders

No quasi-fixed costs

Consider first the simplest albeit unrealistic case characterized by three basic assumptions:

1. There is an infinite number of informed investors endowed with a small amount of the numeraire but no uninformed investor. Each insider can trade only once (with the authority who acts as market maker).

2. There is no quasi-fixed cost of intervention, i.e., $K = 0$.
3. The authority as market maker sets the security price always equal to its beliefs about the fundamental value of the security, i.e., $q = 1 - \left(\frac{\pi_0^R k}{v}\right)^{\frac{1}{2}}$.

Moreover, after every buy (sell) order the authority reduces (increases) its beliefs π_0^R by a small amount so that $\frac{\partial \pi_0^R}{\partial x} < 0$. Consequently, $\frac{\partial q}{\partial x} = -\frac{\partial \pi_0^R}{\partial x} \left(\frac{k}{4\pi_0^R v}\right)^{\frac{1}{2}} > 0$, that is, after the price increases after a buy order and decreases after a sell order.

The authority solves (4.4). Investors anticipate and trade the asset at a closing price reflecting rational expectations (4.7). We will show that in such a setting the following holds:

Proposition 1. *In equilibrium the insiders reveal all their private information. As a result the authority and the informed investors have the same beliefs about the probability of default, that is $\pi_0^R(q^C) = \pi_0$, where q^C is the equilibrium price of the security. This implies*

$$a^I = a^* = \left(\frac{\pi_0 v}{k}\right)^{\frac{1}{2}} - 1$$

and

$$q^C = 1 - \frac{\pi_0}{1 + a^*} = 1 - \left(\frac{\pi_0 k}{v}\right)^{\frac{1}{2}}.$$

The PBE is unique and first best, since the chosen level of action is optimal given π_0 .

Proof. Existence follows directly since the level of protection is optimal and the price reflects the true probability of default, so that no trader is willing to trade the security. Moreover, the beliefs $\frac{\partial \pi_0^R}{\partial x} < 0$ are compatible with the equilibrium: At any price an insider will sell (buy) a security only if its price exceeds (is lower than) its fundamental value (FV), i.e., if $1 - \left(\frac{\pi_0^R k}{v}\right)^{\frac{1}{2}} > (<) 1 - \left(\frac{\pi_0^2 k}{\pi_0^R v}\right)^{\frac{1}{2}}$. This condition holds for $\pi_0 > (<) \pi_0^R$.

In order to prove uniqueness remember first that by sequential rationality the insiders know in advance the beliefs of the authority $\pi_0^{R'}$ and hence the degree of intervention a' , which will be ultimately chosen by the authority at a given price q' . Moreover, in every competitive equilibrium the price of the security for a predicted level of protection must equal the insiders' beliefs about the probability of default, i.e., $q' = 1 - \frac{\pi_0}{1+a'}$ with $a^I = a'$. If this does not hold, there is always an incentive for some insiders to trade the security. In fact, suppose $q' > 1 - \frac{\pi_0}{1+a'}$. Since (4.6) is a smooth function of π_0^R and π_0^R is a smooth function of the order book, an informed investor could then sell a contract (which would decrease q respectively increase a by a small amount) and still make profits. As a result the equilibrium value of q must reveal the beliefs of the informed investors, which the authority knows they represent the true probability of default. Hence, it is not possible that $\pi_0^R(q^C) \neq \pi_0$. Uniqueness follows directly. \square

As an example, assume that nature chooses $\pi_0 > 0.5$. In this case the fundamental value of the contract, which is known by the insiders, is less than the competitive price before the insiders

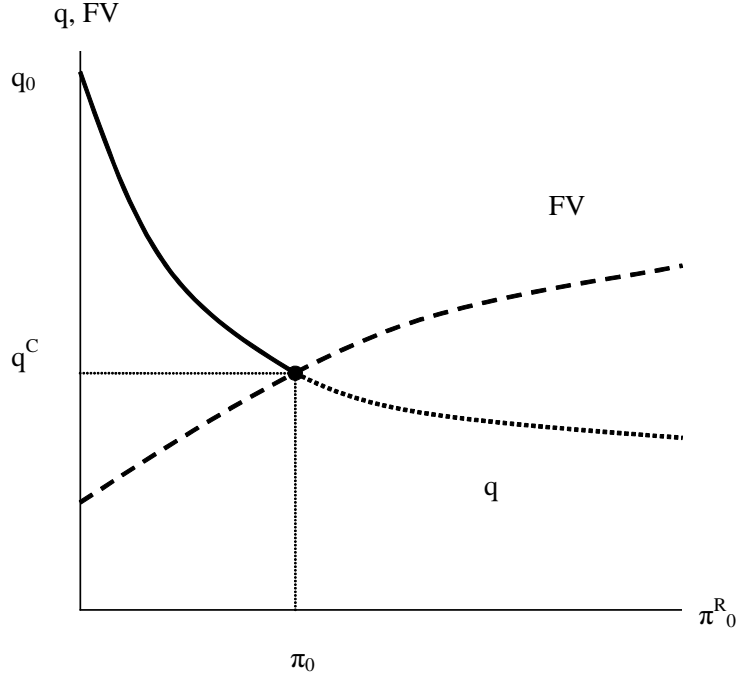


Figure 4.2: Equilibrium without quasi-fixed cost ($K = 0$).

learn π_0 .⁷ As the informed investors start to sell contracts short, q falls, since after each sell order the authority updates its beliefs and π_0^R goes up. At the same time the expectations of the insiders about the degree of intervention ultimately chosen by the authority increase too. As long as $\pi_0^R < \pi_0$ however, some investors want to sell the security and the updating process continues. This will eventually end when the equilibrium values described in Proposition 1 are reached (see Figure 4.2).

Note that the specific functional form of π_0^R does not play any role for the level of the equilibrium price, as long as $\frac{\partial \pi_0^R}{\partial x} < 0$. On the other hand, $q = 1 - \left(\frac{\pi_0^R k}{v}\right)^{\frac{1}{2}}$ is the only pricing process that leads to an equilibrium. To see this, remember that trading stops only if the price equals the fundamental value of the security, that is, if $q = FV = 1 - \left(\frac{\pi_0^2 k}{\pi_0^R v}\right)^{\frac{1}{2}}$. Hence, when insiders cease to trade the authority can infer π_0 from q . However, this presupposes that π_0^R does not change after trading stops. For all π_0 , this is only possible when $q = 1 - \left(\frac{\pi_0^R k}{v}\right)^{\frac{1}{2}}$.

$\tau \left(\frac{\pi_0 k}{v}\right)^{\frac{1}{2}} > \left(\frac{k}{2v}\right)^{\frac{1}{2}}$.

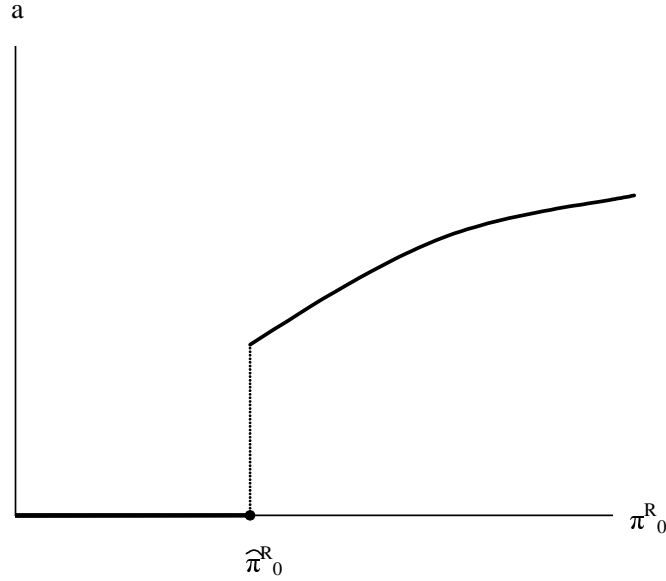


Figure 4.3: Reaction function of the authority with quasi-fixed cost

With a little bit of comparative static, some interesting results may be derived from this simple model. Suppose, for instance, that $k = 0$. In other words, the authority can intervene at no cost and is always successful so that the probability that the bank will ever go bankrupt is zero as long as $\pi_0^R > 0$. Given the assumptions made above $\pi_0^R = 0$ only if $q = 1$. Is this an equilibrium? Yes, but only if the authority intervenes ($a \rightarrow \infty$) also although it thinks that the probability of default is zero. Otherwise no equilibrium exists, since as long as the authority is thought to intervene, the equilibrium price can only be $q = 1$. But if it does not intervene when $q = 1$, the asset (and the bank) becomes risky. This paradox arises in more realistic settings, as shown in the next section.

Quasi-fixed costs

Even if the number of informed investors is infinite, the introduction of quasi-fixed costs has important effects for the equilibrium. In particular, the authority begins to intervene only if its beliefs about the original probability of default are high enough so that the expected gains of intervention cover the quasi-fixed costs. This is the case when $\pi_0^R > \hat{\pi}_0^R$,⁸ as shown in Figure 4.3. As a result the (hypothetical) degree of intervention a would be a non-continuous function of π_0^R . In particular, when the threshold $\hat{\pi}_0^R$ is reached, the next sell order will let the implied degree of intervention jump at the level described by (4.6).

⁸See Appendix 4.6.1 for a derivation of $\hat{\pi}_0$ and $\hat{\pi}_0^R$.

4 The endogeneity pitfall in using market signals as triggers for prompt corrective action

Suppose again that $\frac{\partial \pi_0^R}{\partial x} < 0$ and that the security price always reflects the beliefs of the authority. Then, the authority intervenes only if the price of the security when the market closes is lower than the price implied by $\hat{\pi}_0^R$, which we call \hat{q} . But if the investors are rational, they anticipate that the authority will intervene if the security price goes beyond this specific level. This leads to the following result:

Proposition 2. *In a competitive environment with quasi-fixed costs, the authority begins to intervene only if its beliefs π_0^R are high enough. But when this threshold is reached, the degree of protection jumps at the level described by (4.6). Hence, there is a discontinuity in the “reaction function” of the authority that can considerably reduce the potential for information extraction. As a result, equilibrium may be suboptimal.*

Proof. Obviously, neither do not intervene in any case nor always set a according to (4.6) can be equilibrium strategies. Given $\frac{\partial \pi_0^R}{\partial x} < 0$, the best thing the authority can do is to intervene when π_0^R is high enough, i.e., when the price of the security falls below some value \tilde{q} . In such an equilibrium the competitive price of the security must be equal to $1 - \pi_0$ if the probability of default is less or equal to $1 - \tilde{q}$. Moreover, since (4.5) is a smooth function of π_0^R , at \tilde{q} the authority must be indifferent between intervening or not, i.e., $\pi_0^R = \hat{\pi}_0^R$. This implies that if there is a one-to-one relation between π_0 and the security price, \tilde{q} must be equal to \hat{q} . We will show that this is not the case. In fact, suppose that $\pi_0 = 1 - \hat{q} + \varepsilon$ (ε being a small scalar), that is π_0 is just above the beliefs π_0^R that trigger intervention, and that the price of the security is currently \hat{q} . An informed investor will sell additional securities when

$$\begin{aligned} 1 - \pi(\pi_0, a^I \mid \text{ex trade}) &< \hat{q} \\ 1 - \frac{1 - \hat{q} + \varepsilon}{1 + a_{\text{ex trade}}^I} &< \hat{q} \\ \varepsilon &> a_{\text{ex trade}}^I (1 - \hat{q}) \end{aligned}$$

with

$$a_{\text{ex trade}}^I = \left[\frac{(\pi_0^R \mid \text{ex trade}) v}{k} \right]^{\frac{1}{2}} - 1.$$

But since $a_{\text{ex trade}}^I > 0$, for values of ε small enough the investors have no incentive to signal that $\pi_0 > 1 - \hat{q}$. As a result, only if π_0 exceeds some threshold $\bar{\pi}_0$, the investors would let the price decrease below \hat{q} . The threshold $\bar{\pi}_0$ is the level of π_0 that implies an equilibrium price as defined in Proposition 1 that is not greater than \hat{q} . The problem is then that at $q = \hat{q}$ there would be a jump in the beliefs of the authority since

$$\begin{aligned}
(\pi_0^R \mid q > \hat{q}) &= 1 - q \\
(\pi_0^R \mid q = \hat{q}) &= \frac{(1 - \hat{q}) + \bar{\pi}_0}{2} \\
(\pi_0^R \mid q < \hat{q}) &> \bar{\pi}_0.
\end{aligned}$$

In other words, $(\pi_0^R \mid q = \hat{q})$ is greater than $\hat{\pi}_0^R$. But then it would be unambiguously optimal for the authority to intervene when $q = \hat{q}$. This induces the investors not to let q rise to \hat{q} also when $\pi_0 = 1 - \hat{q}$. But then the authority has to adjust again its beliefs when it observes a price slightly lower than \hat{q} . At the trigger price \tilde{q} the following condition must hold

$$(\pi_0^R \mid q = \tilde{q}) = \frac{(1 - \tilde{q}) + \bar{\pi}_0}{2} = \frac{\pi_0 + \bar{\pi}_0}{2} = \hat{\pi}_0^R.$$

Hence, for $\pi_0 \in [\underline{\pi}_0, \bar{\pi}_0]$ the last trade takes place slightly above \tilde{q} . Therefore, there is no one to one relationship between π_0 and the security price, which may be a very imprecise signal of the true value of π_0 (see Figure 4.4). For $\pi_0 \leq \hat{\pi}_0$ this has no welfare implications as $a > 0$ would not be efficient. But for $\pi_0 \in [\hat{\pi}_0, \bar{\pi}_0]$ the equilibrium is not first best in terms of welfare, as optimal intervention would cover the quasi-fixed costs for all $\pi_0 > \hat{\pi}_0$.

The beliefs $\frac{\partial \pi_0^R}{\partial x} < 0$ are again compatible with the equilibrium. As opposed to the setting without quasi-fixed costs, π_0^R is not a smooth function of the trading volume, though. For instance, when an insider sells a security at $q = \hat{q}$, the beliefs of the authority increase from π_0^R to above $\bar{\pi}_0$. \square

Clearly, for a subset of π_0 a fully revealing equilibrium exists. Suppose, for instance, that $\pi_0 > \bar{\pi}_0$. Then, the insiders have the incentive to sell the security also at the price that triggers supervisory intervention, since they know that the final price/FV of the security will be $q^C < \hat{q}$ (see Figure 4.5). Note that the security price actually never equals \tilde{q} . After the sell order just above \tilde{q} the price jumps to \hat{q} since $\pi_0^R = \hat{\pi}_0^R$. Instead, when an insider sells at \hat{q} , the beliefs of the authority jumps from $\hat{\pi}_0^R$ to $\bar{\pi}_0$ but the price of the security decreases only slightly.

The higher the difference between π_0 and $\bar{\pi}_0$, the more likely the price will be an imprecise signal of the true value of π_0 . In particular, this is the case when the quasi-fixed cost component K in the loss function of the authority is relatively large.

Again, the case with $k = 0$, is particularly interesting. The basic paradox described in Section 4.2.1 is very similar to the case where the authority faces only quasi-fixed costs. In such a case, if the price of the contract falls below some threshold, the authority reduces the probability of default to zero, so that also the value of the security rises to one. Note that in this case $\bar{\pi}_0 = 1$ so that no insider trades at q_0 if (4.5) is (not) satisfied at the outset of the game and $\pi_0 < (>) \frac{1}{2}$. The case of pure strategies with $k = 0$ may be interpreted as a case in which the authority has a particular rule for intervention (e.g., it intervenes if the spread on subordinated debt exceeds 3%) and when it intervenes, it is always successful.

What about semi-separating equilibria and mixed strategies? The authority may randomize only when indifferent between intervening or not, that is when $\pi_0^R = \hat{\pi}_0^R$. For all other beliefs there is

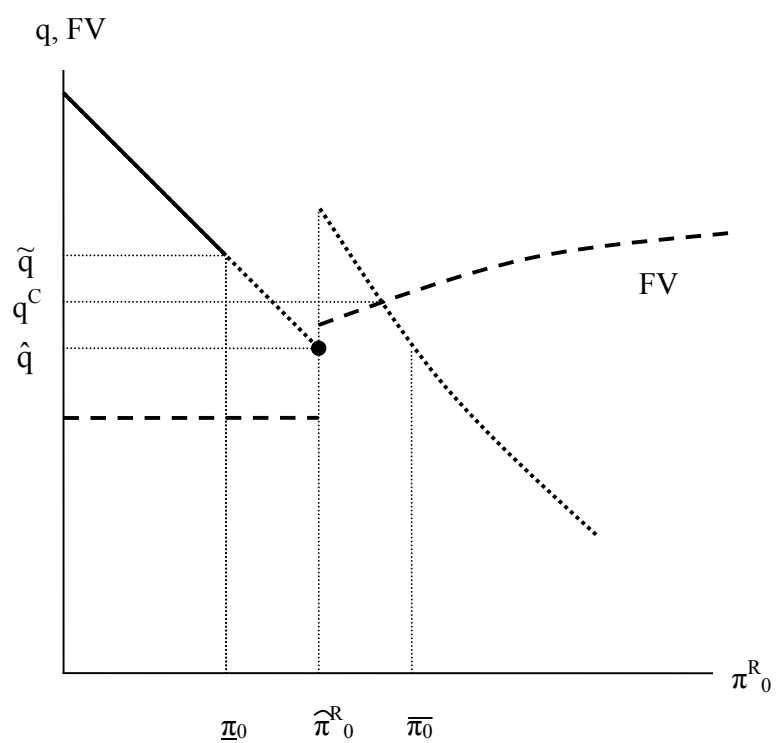


Figure 4.4: Partially revealing equilibrium with quasi-fixed cost.

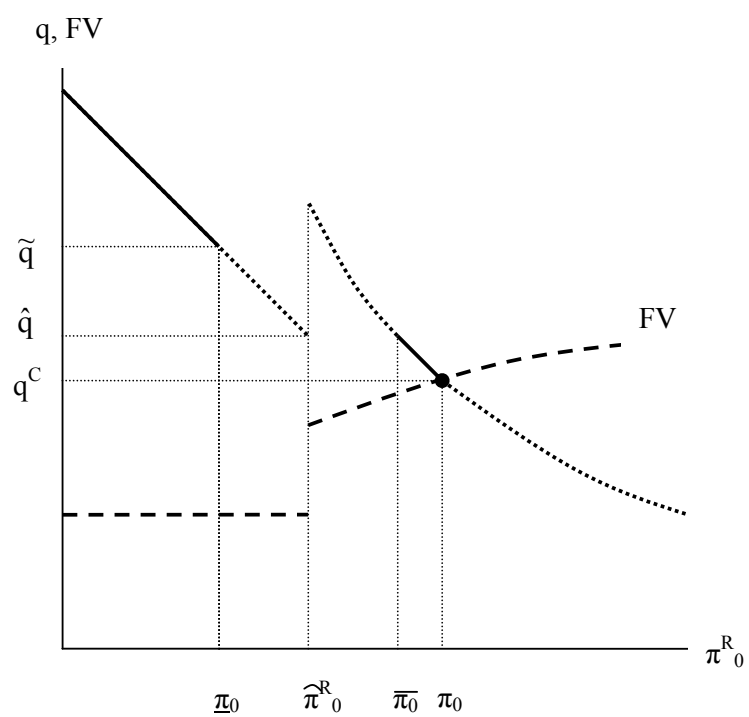


Figure 4.5: Fully revealing equilibrium with quasi-fixed cost.

4 The endogeneity pitfall in using market signals as triggers for prompt corrective action

a single optimal action. As long as the probability of $a > 0$ is strictly greater than zero at $\hat{\pi}_0^R$, the investors will never let the price decrease to \tilde{q} when $\pi_0 \in [\underline{\pi}_0, \bar{\pi}_0]$, so that in this case the equilibrium price equals $\tilde{q} - \varepsilon$.

Noise traders

Let us go back to the setting with $K = 0$. Clearly, the assumption that only informed traders enter the market is very unrealistic. In fact, speculative markets in which a large amount of investors possess clear information advantages usually do not set up in reality. No uninformed speculative investor respectively market maker is willing to trade in such a market, since they know they can only lose money. Above market liquidity was supplied by the authority, however under the strong assumption that its loss function does not include trading income.

As noted in Section 4.3.1, Kyle (1985) shows that with a sufficient amount of noise trading it is possible to overcome this adverse selection problem.

We model noise trading in a very simple way:

$$\begin{aligned}\Pr(x > 0 \mid \text{noise trade}) &< \Pr(x < 0 \mid \text{noise trade}) \text{ if } q > q_0 \\ \Pr(x > 0 \mid \text{noise trade}) &> \Pr(x < 0 \mid \text{noise trade}) \text{ if } q < q_0\end{aligned}$$

that is, if the current price of the security is more (less) than the equilibrium price without asymmetric information, the probability of a buy order from a noise trader is less (more) than the probability of a sell order. In other words, the noise trader is more inclined to buy the security if the current price of the security is less than the equilibrium price without asymmetric information. One can interpret noise traders as naive “insiders” that have uniformly distributed beliefs about π_0 and do not use public information q in their trading strategy. We assume, in addition, that at every trade the market maker and the authority do not know whether they face an informed or a noise investor.

Under those assumptions the risk-neutral insiders will trade the security until the price reflect the intrinsic value of the contract, as in Section 4.3.4. This is because the insider knows that at this price the probability of an unfavorable trade – a noise sell (buy) order if the insider just purchased (sold) the security, which decreases (increases) a – is less than the probability of a favorable trade.⁹ Moreover when noise trades shift the price away from its equilibrium value, the informed investors ensure that the price goes back to q^C . Hence, Proposition 1 applies.¹⁰

The introduction of noise traders can even lead to the existence of a fully revealing equilibrium with quasi-fixed costs for all values of π_0 . This is because there is always a non-zero probability of

⁹Note that since $\frac{\partial^2 a}{\partial (\hat{\pi}_0^R)^2} < 0$, this is always true only when the effect of each trade on the price (and hence on a) is small. If $\frac{\partial a}{\partial x}$ is large, the difference between $\Pr(x > 0 \mid \text{noise trade})$ and $\Pr(x < 0 \mid \text{noise trade})$ must high enough, so to compensate for the convexity of the reaction function of the supervisor.

¹⁰If the last trade before market closing is a noise trade the equilibrium values described in Proposition 1 hold only approximately.

a noise trade at the price that triggers supervisory intervention. Hence, insiders can trade until this price and then wait until a noise trader triggers supervisory intervention (this may require several attempts, as the direction of the next noise trade is unknown). Note that since the supervisor does not know whether it faces an insider or a noise trader, it cannot be sure that the trade at the trigger price is a noise trade, which would imply $\frac{\partial \pi_0^R}{\partial x} = 0$. In fact, remember that for some values of π_0 , insiders would also trade at the trigger price. As a corollary noise traders cannot realize when trading at the trigger price leads to a sure loss and hence assuming a non-zero probability of a noise trade (buy or sell order) is not inconsistent.

4.3.5 Equilibrium: Asymmetric information and monopolistic insiders

No quasi-fixed costs

Up until now, we have considered a framework where an infinite number of insiders with a small amount of the numeraire enter the market. In practice, however, few (if any) investors possess clear superior information. In order to see the consequences of relaxing the competition assumption in our model, suppose that there is only one informed investor endowed with an infinite amount of the numeraire.

We saw in Section 4.3.2 that when a monopolistic insider begins to accumulate inventories i , she stops trading before the price reaches the fundamental value of the security. More specifically, the informed investor stops trading when

$$q^S = 1 - \pi - i \frac{\partial \pi}{\partial x}. \quad (4.8)$$

This implies the following:

Lemma 3. *A first best, fully revealing equilibrium under which the price of the security always reflects the beliefs of the authority/market maker and $\frac{\partial \pi_0^R}{\partial x} < 0$ cannot exist.*

Proof. Such a fully revealing equilibrium implies $\pi_0^R = \pi_0$ and $a = \left(\frac{\pi_0 v}{k}\right)^{\frac{1}{2}} - 1$. But since $\frac{\partial \pi}{\partial x} > 0$ by the assumptions $\frac{\partial \pi_0^R}{\partial x} < 0$, (4.8) can never be satisfied for $q = 1 - \left(\frac{\pi_0 k}{v}\right)^{\frac{1}{2}} = 1 - \pi$ and $i \neq 0$. \square

Note that this distortion exists only because the fundamental value of the security is endogenous, i.e., it depends on the action taken by the supervisor, which, in its turn, depends on the price of the security. In a traditional market $\frac{\partial \pi}{\partial x} = 0$ and this feedback effect is not present.

This effect of inventories on trading incentives exists also in the presence of noise traders and other insiders. Hence, if the supervisor always sets q according to his beliefs, in a market with of a finite number of informed investors it will never observe a price that reflects the true probability of default. Moreover, if an equilibrium exists, it must be that $q^S > q^C$ for $i < 0$ and $q^S < q^C$ for $i > 0$.

We pointed out above that since trading income is not included in the loss function of the supervisor, there is in principle no restriction on how the supervisor should adjust the security

4 The endogeneity pitfall in using market signals as triggers for prompt corrective action

price after each trade. Moreover, with only one informed trader the supervisor knows that the (net) amount of outstanding contracts equals the (net) portfolio of the insider. As a result, there is a one-to-one relationship between π_0 and the security price. This implies the following:

Proposition 4. *With a known number of informed investors N there exists an equilibrium so that $\pi_0^R(q^S) = \pi_0$ and a is optimal given π_0 . Hence, the equilibrium is first best. The equilibrium price q^S is given by¹¹*

$$q^S = 1 - \left(\frac{\pi_0 k}{v} \right)^{\frac{1}{2}} + \frac{I}{N} \frac{\partial \pi_0(I)}{\partial x} \left(\frac{k}{4\pi_0 v} \right)^{\frac{1}{2}},$$

where the sum of the long and short positions of all insiders I equals $\sum_{x \in X} x$, the net amount of securities traded (buy minus sell orders). Each insider n has the same inventories $i_n = \frac{I}{N}$.

Proof. With a little algebra, it is straightforward to show that q^S satisfies (4.8) for each insider. Moreover, since the supervisor knows I , it can infer the true value of π_0 and optimally choose a , so that for the supervisor as well there is no reason to deviate from equilibrium. By the same argument outlined in the competitive case, the only pricing schedule that can achieve this equilibrium is one with

$$q = 1 - \left(\frac{\pi_0^R k}{v} \right)^{\frac{1}{2}} + \frac{I}{N} \frac{\partial \pi_0^R(I)}{\partial x} \left(\frac{k}{4\pi_0^R v} \right)^{\frac{1}{2}}. \quad (4.9)$$

Since $\frac{\partial \pi_0^R}{\partial x} < 0$ the price of the contract is lower (higher) than in the competitive case if $\pi_0 < (>) 0.5$. When the market opens $I = 0$ and $q_0 = 1 - \left(\frac{k}{2v} \right)^{\frac{1}{2}}$ as in the competitive case. Hence, the insiders start buying (selling) and accumulates long (short) positions only if $\pi_0 < (>) 0.5$. Hence, the beliefs $\frac{\partial \pi_0^R}{\partial x} < 0$ are compatible with the equilibrium. \square

The equilibrium price with monopolistic insiders differs from the fundamental value of the security and it is not efficient in the semi-strong sense. Moreover, also all along the equilibrium path q is lower (higher) than the expected value of the security for $\pi_0 < (>) 0.5$. This implies that the trading losses of the supervisor as market maker are higher than in the competitive case. In order to extract more information, the supervisor must give the monopolist additional rent.

As opposed to Proposition 1, how the beliefs of the authority react to each trade directly affects the equilibrium price.

If the supervisor does not know how many insiders it faces and their portfolios, the relationship between q and π_0 is not unique anymore, but depends on the actual value of N . In this case the supervisor has to build expectations \bar{N} about the number of insiders. By posting the supply schedule (4.9) as if $N = \bar{N}$, when $N > \bar{N}$ the price of the security will overshoot the fully revealing price, as q^S increases with N . This implies that for high values of π_0 the supervisor will intervene too strongly. When $N < \bar{N}$, the opposite is true.

¹¹For the derivation of q^S see Appendix 4.6.2.

Quasi-fixed costs

Also in the monopolistic case the introduction of quasi-fixed costs has several consequences. While the FOC (4.8) of the insider remains the same, as long as $q < \tilde{q}$ the supervisor does not intervene and hence $\frac{\partial \pi}{\partial x} = 0$. In other words, the optimization problem of the monopolist is equivalent to the one of the competitive insider. Above we saw that for $\pi_0 > \bar{\pi}_0$ the competitive insider is willing to sell at \hat{q} , which triggers supervisory intervention. In the monopolistic case intervention increases the value of all contracts already sold. Hence, $\bar{\pi}_0$ is higher and \tilde{q} lower than in the competitive case. Intuitively, an insider with a short position has even less incentives than a competitive insider to let the price decrease below the trigger price, since this would significantly raise the value of all the securities she already sold. As a result, the range of π_0 for which the equilibrium is rather uninformative is negatively correlated with N .

Noise trading

The introduction of noise traders has important effects on the equilibrium as well. First, we saw above that the supervisor needs to know the portfolio of the insiders in order to derive π_0 from the security price. In the presence of noise traders this may not be possible, even if N were common knowledge. Hence, the supervisor has to build expectations on the probability of a noise respectively informed trade and to adjust π_0^R and q accordingly.

Second, as time passes and the insiders accumulate inventories, the price becomes more and more noisy. To see this, suppose that we are in the situation where (4.8) is satisfied and $\pi_0 > \frac{1}{2}$, which implies that in all probability the insider has a short position in the security. Given our assumption, the next trade is likely to be a buy order from a noise trader. Given $\frac{\partial \pi_0^R}{\partial x} < 0$,¹² the price would then decrease and the insider has again an incentive to sell an additional contract. If the next trade is again a buy order from a noise trader, however, this may not be the case anymore, since $|i|$ has increased by an additional unit. Hence, in the long run, the price tends to go back to its unconditional mean $\frac{1}{2}$. As a result, in the presence of monopolistic insiders and noise traders, the supervisor should not leave the market open for too long, since as time passes the security price is likely to become less and less informative. Yet, a rational supervisor should realize that as time passes noise trades become more and more likely and hence it should reduce $\left| \frac{\partial \pi_0^R}{\partial x} \right|$ accordingly. In the real world, however, such adjustments are very difficult since they require a high amount of information.

¹²Note that since noise traders can trade in both directions at any time, the supervisor can never know who is trading at a particular price. Hence, $\frac{\partial \pi_0^R}{\partial x} < 0$ is always justified.

4.4 Discussion

4.4.1 The mechanics behind the endogeneity problem

We modeled the strategic interaction between an insider who knows the ex ante probability of an event and an authority who would like to prevent the event from happening. The interaction is implemented by a market for a contract written on the event. The authority offers the contract, insiders can buy. The authority observes the price and/or the order book, infers the ex ante probability of the event, and intervenes to prevent it. This is a game with private information by the insiders and hidden action by the authority.

We have found that the market breaks down, if government intervention is costless and always successful to one hundred per cent. Any deviation of the asset price from unity would lead to full prevention of the event (“zero tolerance”). More surprisingly, we have found that if intervention is costly (i.e., if marginal prevention-cost is positive but the quasi-fixed cost is zero), not only does the market not break down: The equilibrium is even first best with full revelation of private information. There is an equilibrium in which insiders correctly predict the authority’s action (the level of preventive effort) and the authority correctly interprets the contract price paid by investors.

If the authority dislikes intervention per se (modeled as a quasi-fixed cost), an equilibrium may or may not exist. Even if an equilibrium exists the authority may choose an inefficient intervention level. The results from the quasi-fixed cost model seem to contradict the well known result (see, e.g., Brunnermeier, 2001) that in this type of model private information is revealed in prices if everything else is common knowledge. However, that result rests on the assumption of complete markets. In our model markets are not complete in the sense that there is no asset written on the authority’s action.¹³ In the absence of a quasi-fixed cost this is not a problem, because the action can be inferred. Under a quasi-fixed cost this is no longer true, and market incompleteness takes its toll.

The assumption of a quasi-fixed cost seems quite plausible not only a priori, but also in the light of our result. We find that in the presence of quasi-fixed cost, the market price of an asset written on bank solvency is not a monotonous function of the prior (pre-intervention) probability of failure. This helps understand the otherwise “perverse” reaction of bank debt prices reported in DeYoung, Flannery, Lang, and Sorescu (2001), namely an *increase* after unexpectedly bad examination results (suggesting that the intervention threshold was passed).

The results from the quasi-fixed cost model also lend a formal underpinning to the intuition formulated in Bond, Goldstein, and Prescott (2006): “The inference from the price is non-trivial. For example, a low price may indicate that the fundamentals are bad, and thus call for the authority’s intervention. It may also indicate that fundamentals are not bad enough to justify intervention in which case the price is low because no intervention is expected” (Bond, Goldstein, and Prescott, 2006, p. 5f.).

¹³This is why in Bond, Goldstein, and Prescott (2006) the introduction of a “regulator security” leads to unique revealing equilibrium.

All results from our model depend on one crucial assumption: that players rationally guess each others actions and that rationality is common knowledge. As is well known in game theory, out-of-equilibrium beliefs are hard to model. If one assumes that one party behaves irrationally or anticipates another party to do so, not much can be said about the outcome without some further assumptions.

A crucial issue not treated in the paper in a comprehensive manner is noise. We assumed that the authority is a “perfect” market maker, setting a bid-ask spread of zero. In a more realistic setting, where trade is made possible by noise traders, prices would become noisy, making prediction difficult. In general, market prices are not very informative in areas where the noise in prices is high relative to the level of risk. A striking example is the airline industry. In theory one would expect passengers to be ready to pay a bit more for a ticket on a safe flight than on a risky trip. Yet, airfares, even on the same flight, differ so widely relative to the very low probability of an accident. Therefore, nobody has ever proposed to read the risk of a crash of different companies or of different aircraft from observed airfares.

4.4.2 A way around endogeneity: Price neutral intervention

Our model rests on one assumption we have not explicitly discussed yet: The existence of one single asset paying one dollar if the bank remains solvent and nothing if it fails. We will now relax this assumption in order to show a way around the double endogeneity effect. The idea is to construct an asset with an expected payoff that depends only on the bank’s pre-intervention solvency but not on the degree of supervisory intervention.¹⁴

Think of an asset that pays nothing if the bank defaults (if $\omega = 0$). If the bank remains solvent ($\omega = 1$) the asset pays $h(a)$, where h stands for “haircut”, with $h(0) = 1$ and $\frac{\partial h}{\partial a} < 0$. In words: If the authority intervenes ($a > 0$), the nominal value of the asset is reduced to a fraction h . The stronger the intervention, the stronger the reduction or the “shorter” the haircut.

The asset is insulated from the effect of supervisory intervention if the following condition holds: The increase in real value (expected payoff) caused by an intervention must exactly be compensated by the concomitant reduction in its nominal value. This yields the following condition for the “neutral” haircut function $h(a)$:

$$\left(1 - \frac{\pi_0}{1 + a}\right) h(a) = 1 - \pi_0. \quad (4.10)$$

The left-hand side is the expected return on the asset *with* an intervention a . The right-hand side is the expected return *without* any intervention. If h satisfies the equality between both sides the expected payoff of the asset does not react to intervention; it is always $1 - \pi_0$. From the market value of the asset (still assuming there is a competitive market), the authority can directly read the prior probability of failure π_0 .

¹⁴We read the proposal in Bond, Goldstein, and Prescott (2006) to link intervention with a tax on those who benefit from intervention as a binary action space version of the proposal made in the text.

4 The endogeneity pitfall in using market signals as triggers for prompt corrective action

Solving (4.10) for $h(a)$ yields:

$$h(a) = \frac{(1+a)(1-\pi_0)}{1+a-\pi_0}. \quad (4.11)$$

This is the price-neutral haircut function. It assures that the competitive price of the asset q always reflects the original failure probability, $q = 1 - \pi_0$. The asset with payoffs $\{0, h\}$ thus is the ideal indicator asset. Its price reflects the prior default probability even if it is common knowledge that the authority will intervene and that the posterior default probability will differ from the prior probability.

Two remarks apply, one theoretical, one practical. From a theory point of view, (4.11) seems to violate our assumptions: It implies an asset indirectly written on the prior probability of failure π_0 which is unobservable and thus not contractible. However, it is straightforward to show¹⁵ that the haircut function (4.11) based on the beliefs of the authority π_0^R together with the competitive market provide a truth-telling mechanism. An insider who learns π_0 has an incentive to reveal π_0 directly by trading at a price $q = 1 - \pi_0^R$.

The intervention-neutral asset may seem a theoretical artefact without any resemblance in reality. Yet, there exist arrangements with a striking similarity to such an asset. For example the 1991 FDICIA includes a provision authorising the FDIC to settle uninsured and unsecured claims on an institution in receivership with a final settlement payment which must reflect an average of the FDIC's receivership recovery experience.¹⁶ Such a settlement would imply a haircut to creditors' nominal claims reflecting expected losses. It tends to insulate the market value of these claims from the intervention by the receiver (the FDIC). Although safeguarding the informational value of debt prices was probably not the legislator's purpose, the provisions seem to work in that direction.

In practice, a security subject to a price-neutral haircut function may hardly be implementable via private contracts alone. Rather, the intervention-contingent haircut would have to be implemented by a combination of private contracts and rules on bank insolvency or restructuring, as in the FDICIA example. The following set of rules for the authority would seem to do the job:

1. Define an "indicator security" (e.g., subordinated debt).
2. Declare in advance that in case of intervention the nominal value of the indicator security will be cut to a fraction h of the original nominal value.
3. Set h to the pre-intervention fundamental value of the indicator security in percentage of the fundamental value of the indicator security without haircut.

These rules implement the price-neutral haircut function. It is straightforward to verify that the (competitive) market price of the indicator security, or more precisely its risk spread, is a direct and unbiased measure of the bank's default probability as seen by the market insiders.¹⁷

¹⁵See Appendix 4.6.3.

¹⁶FDICIA, Title IV, Subtitle B.

¹⁷The legal applicability of such rules depends on the particular legislation. A necessary condition is the existence of

4.5 Conclusions

Can – and should – bank supervisors base prompt corrective action on market data? The literature on the information content of market data, like prices for bank debt and equity, would at least partially suggest so. Yet, we have found that the right answer is mixed.

To begin with, market prices have to be used with caution. Not because market data are too noisy or because markets are sometimes wrong, but because the expected use of market data feeds back on these data. The simple example is a high subordinated debt spread which triggers a successful supervisory intervention and thus belies itself. The spread becomes a “self-defeating prophecy”. This is the double endogeneity effect of an action which is (i) price-dependent and (ii) rationally predicted (and priced). It seems that this effect has been largely overlooked in a supervisory context, although inspiration could be taken from the literature on exchange rates or on central banks and asset bubbles.

The double endogeneity effect may not invalidate prices as guides for action in all cases. A job candidate may be right to reject a job offer in the light of a recent sharp drop in the share price of the company (Allen, 1993). The impact of the candidate on the value of the company is small. It is unlikely that the fall in the share price reflects market anticipations that the candidate will reject the job although she is the only person who could save the company.

An authority, by contrast, is not atomistically small. Its decisions, unlike those of a job candidate, have a measurable influence on the fundamentals on which prices are based. If market prices incorporate really all information available to the market, this includes information or beliefs about what the supervisor is going to do. Market prices for bank debt, for example, if used as a basis for supervisory decisions, become subject to the endogeneity problem.

Still this does not invalidate market prices. We have shown that when the authority is “predictable” (when its reaction function is common knowledge), an equilibrium may still exist. However, market prices in such an equilibrium have to be read correctly: They do not reflect the probability of a bank failure to which the authority should perhaps react, but rather the probability of bank failure *after* the authority has taken the action expected by the market. A failure to recognize the double endogeneity effect could lead authorities astray.

We have also shown that an equilibrium is less likely to exist when the authority dislikes to exceed some particular level of intervention. We used a quasi-fixed cost of any action exceeding a level of zero. But this level could arbitrarily be set. A realistic assumption is that the cost function has an upward jump at the point where an intervention becomes public.

Advocates of market based intervention have stressed the disciplining effect on the supervisory authority of a price-based rule, for example of a subordinated debt spread policy. Our findings suggest that in that very case, when a rule would be most needed because the authority has a

some bank liability – like subordinated debt or preferred equity – that can be reduced by the supervisor without liquidation of a bank. The Swiss banking act, for example, empowers the supervisory authority to reduce non-protected creditors’ claims as part of mandatory restructuring (Swiss Banking Act, Art. 29). It does not require such a haircut to be “price-neutral” in the above-defined sense, though.

4 The endogeneity pitfall in using market signals as triggers for prompt corrective action

dislike to intervene (or to transcend a certain level), an equilibrium may not exist, or the market may fail entirely.

Our message that market prices may lose their information content if authorities try to use them for preventive purposes is applicable to other fields as well. For example, the claim that the market for “terrorism futures” proposed by an US Ministry of Defense agency in July 2003, was “a good idea with bad press” (Hal Varian) would only be supported by our model under very restrictive assumptions. These assumptions are unlikely to hold in reality.

There are several assumptions needed for an equilibrium with revealing prices that may not be met in reality, like: (i) sufficiently low noise, (ii) common knowledge of preferences (the authority’s prevention cost function and individuals’ attitudes towards risk), (iii) perfect competition among informed individuals. We have to leave it to further research to clarify the impact of factors like noise in prices on the double endogeneity effect.

Reality is not only a source of complications, but also a source of hope. The double endogeneity circle can be broken if supervisors find ways to intervene in a bank in a “neutral” way, that is without affecting the payoffs to the holders of those very assets whose prices are used as market signals. If a supervisor looks at subordinated debt spreads, for example, then neither a closure of a bank (which would mean a loss on subordinated debt) nor a bailout (which would extend to subordinated debt holders) would satisfy the neutrality criterion. The design of appropriate rules is a line of further practical research.

4.6 Appendix

4.6.1 The trigger value of π_0

In a model with quasi-fixed costs, it is efficient for the authority to intervene only if π_0 exceeds some value $\hat{\pi}_0$. This trigger value $\hat{\pi}_0$ solves

$$K = \hat{\pi}_0 v - \frac{\hat{\pi}_0 v}{1+a} - ka$$

with

$$a = \left(\frac{\hat{\pi}_0 v}{k} \right)^{\frac{1}{2}} - 1. \quad (4.12)$$

The first equation reduces to

$$\begin{aligned} K &= \hat{\pi}_0 v \left(\frac{a}{1+a} \right) - ka \\ &= a \left(\frac{\hat{\pi}_0 v}{1+a} - k \right) \end{aligned}$$

Inserting (4.12)

$$K = \left[\left(\frac{\widehat{\pi}_0 v}{k} \right)^{\frac{1}{2}} - 1 \right] \left[(k \widehat{\pi}_0 v)^{\frac{1}{2}} - k \right]$$

so that

$$\widehat{\pi}_0 = \frac{k + 2(kK)^{\frac{1}{2}} + K}{v}.$$

Accordingly, the authority intervenes only when

$$\widehat{\pi}_0^R > \frac{k + 2(kK)^{\frac{1}{2}} + K}{v}.$$

As one would expect, this condition is more likely to be satisfied when the costs of intervention (variable k and quasi-fixed K) are low and the gains (v) are high.

4.6.2 Equilibrium price with monopolistic insiders

If there is only one insider trading the security the equilibrium must satisfy the following condition

$$q^S = 1 - \pi - i \frac{\partial \pi}{\partial x}$$

where i is the net inventory (long minus short positions) of the monopolist and

$$\frac{\partial \pi}{\partial x} = \frac{\partial \pi}{\partial a} \frac{\partial a}{\partial \pi_0^R} \frac{\partial \pi_0^R}{\partial x}.$$

Given our specifications, $\frac{\partial \pi}{\partial a} = -\frac{\pi_0 k}{\pi_0^R v} < 0$ and $\frac{\partial a}{\partial \pi_0^R} = \left(\frac{v}{4\pi_0^R k} \right)^{\frac{1}{2}} > 0$ if (4.5) is satisfied. Moreover, in the presence of only one informed trader $\mathbf{M} = \{i\}$ as the insider trades only in one direction (either buy or sell order only). Consequently, π_0^R depends only on i and

$$q^S = 1 - \pi - i \left(-\frac{\pi_0 k}{\pi_0^R v} \right) \left(\frac{v}{4\pi_0^R k} \right)^{\frac{1}{2}} \frac{\partial \pi_0^R(i)}{\partial x}.$$

In a fully revealing equilibrium $\pi_0 = \pi_0^R$, so that

$$\begin{aligned} q^S &= 1 - \pi + i \frac{\partial \pi_0(i)}{\partial x} \left(\frac{k}{4\pi_0 v} \right)^{\frac{1}{2}} \\ q^S &= q^C + i \frac{\partial \pi_0(i)}{\partial x} \left(\frac{k}{4\pi_0 v} \right)^{\frac{1}{2}} \end{aligned}$$

Since $\frac{\partial \pi_0^R(i)}{\partial x} < 0$, the monopolistic fully revealing price must be higher than the competitive price when $i < 0$ and vice versa.

4 The endogeneity pitfall in using market signals as triggers for prompt corrective action

With N insiders the net amount of contracts traded equals the sum of the positions of each insider, that is $\sum_{x \in X} x = \sum_{n=1}^N i_n = I$, where i_n is the net portfolio of insider n . In the oligopolistic case trading stops when

$$q^S = 1 - \pi - i_n \frac{\partial \pi}{\partial x} \quad \text{for all } n.$$

Since all the insiders have the same information set, all N conditions are simultaneously satisfied only when $i_1 = i_2 = \dots = i_N = \frac{I}{N}$. As a result, if the supervisor knows N it can infer π_0 from the fully revealing equilibrium price

$$q^S = q^C + \frac{I}{N} \frac{\partial \pi_0(I)}{\partial x} \left(\frac{k}{4\pi_0 v} \right)^{\frac{1}{2}}$$

since all other determinants of q^S are known. Note that $q^S \rightarrow q^C$ for $N \rightarrow \infty$.

4.6.3 Equilibrium with price neutral intervention

Suppose that the asset traded has an expected payoff that depends on the degree of supervisory intervention, that is it pays nothing if the bank defaults and pays $h(a)$ otherwise. $h(a)$ is set by the authority as follows:

$$h(a) = \frac{(1+a)(1-\pi_0^R)}{1+a-\pi_0^R}.$$

When the authority as market maker sets the price of the security q equal to its beliefs about the fundamental value, we have

$$\begin{aligned} q &= (1-\pi^R) h(a) \\ &= 1 - \pi_0^R. \end{aligned}$$

Note that the price of the security is a smooth function of π_0^R even in the presence of quasi-fixed costs.

For the investor the fundamental value of the security is given by

$$\begin{aligned} FV &= (1-\pi) h(a) \\ &= (1-\pi_0^R) \left(\frac{1+a-\pi_0}{1+a-\pi_0^R} \right). \end{aligned}$$

In a competitive setting investors trade the security as long as its price differs from its fundamental value. For instance the investors buy securities when $FV > q$. But $FV > q$ only when $\pi_0 < \pi_0^R$. Given the beliefs $\frac{\partial \pi_0^R}{\partial x} < 0$, we can have a unique competitive equilibrium with $q = FV$ only if $\pi_0^R = \pi_0$, i.e., the equilibrium is fully revealing.

Bibliography

- ALLEN, F. (1993): “Stock Markets and Resource Allocation,” in *Capital Markets and Financial Intermediation*, ed. by C. Mayer, and X. Vives, pp. 81–116. Cambridge University Press, Cambridge UK.
- ALTUNBAS, Y., L. EVANS, AND P. MOLYNEUX (2001): “Bank Ownership and Efficiency,” *Journal of Money, Credit, and Banking*, 33(4), 926–954.
- ANDREWS, A. M. (2005): “State-owned Banks, Stability, Privatization, and Growth: Practical Policy Decisions in a World Without Empirical Proof,” IMF Working Paper 05/10, International Monetary Fund.
- ARORA, N., J. R. BOHN, AND F. ZHU (2005): “Reduced Form vs. Structural Models of Credit Risk: A Case Study of Three Models,” *Journal of Investment Management*, 3(4), 43–67.
- BASEL COMMITTEE ON BANKING SUPERVISION (2003): “Markets for Bank Subordinated Debt and Equity in Basel Committee Member Countries,” BCBS Publications No. 12.
- BERGER, A. N., S. M. DAVIES, AND M. J. FLANNERY (2000): “Comparing Market and Supervisory Assessments of Bank Performance: Who knows what when?,” *Journal of Money, Credit, and Banking*, 32(3), 641–667.
- BIAIS, B., D. MARTIMORT, AND J.-C. ROCHET (2000): “Competition Mechanisms in a Common Value Environment,” *Econometrica*, 68(4), 799–837.
- BICHSEL, R. (2006): “State-Owned Banks as Competition Enhancers, or the Grand Illusion,” *Journal of Financial Services Research*, 30(2), 117–150.
- BIRCHLER, U. W., AND M. FACCHINETTI (2007): “Can Bank Supervisors Rely on Market Data? A Critical Assessment from a Swiss Perspective,” *Swiss Journal of Economics and Statistics*, 143(2), 95–132.
- BIRCHLER, U. W., AND D. HANCOCK (2004): “What Does the Yield on Subordinated Bank Debt Measure?,” Finance and Economics Discussion Series 2004-19, The Federal Reserve Board.
- BIRCHLER, U. W., AND M. A. MAECHLER (2002): “Is there Market Discipline in Swiss banks?,” *Research in Financial Services, Prompt Corrective Action in Banking: 10 Years later*, 14, 243–257.

Bibliography

- BLACK, F., AND J. C. COX (1976): "Valuing Corporate Securities: Some Effects of Bond Indenture Provisions," *The Journal of Finance*, 31(2), 351–367.
- BLACK, F., AND M. S. SCHOLES (1973): "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81(3), 637–54.
- BLISS, R. R., AND M. J. FLANNERY (2002): "Market Discipline in the Governance of US Bank Holding Companies: Monitoring vs. Influencing," *European Finance Review*, 6(3), 361–395.
- BOARD OF GOVERNORS OF THE FEDERAL RESERVE SYSTEM (1999): "Using Subordinated Debt as an Instrument of Market Discipline," Staff Study 172.
- BOND, P., I. GOLDSTEIN, AND E. S. PRESCOTT (2006): "Market-Based Regulation and the Informational Content of Prices," Federal Reserve Bank of Richmond Working Paper 06-12, University of Pennsylvania and Federal Reserve Bank of Richmond.
- BONGINI, P., L. LAEVEN, AND G. MAJNONI (2002): "How good is the market at assessing bank fragility? A horse race between different indicators," *Journal of Banking and Finance*, 26(5), 1011–1028.
- BREWER, E., H. GENAY, W. C. HUNTER, AND G. KAUFMAN (2003): "Does the Japanese Stock Market Price Bank Risk? Evidence from Financial Firm Failures," *Journal of Money, Credit, and Banking*, 35(4), 507–543.
- BRUNNERMEIER, M. M. (2001): *Asset Pricing under Asymmetric Information*. Oxford University Press, Oxford U.K.
- BUSER, S. A., A. H. CHEN, AND E. J. KANE (1981): "Federal Deposit Insurance, Regulatory Policy, and Optimal Bank Capital," *The Journal of Finance*, 36(1), 51–60.
- CALOMIRIS, C. W. (1999): "Market-Based Banking Supervision," *The Financial Regulator*, 4, 33–36.
- CAMPBELL, J. Y., AND G. B. TAKSLER (2003): "Equity Volatility and Corporate Bond Yields," *The Journal of Finance*, 58(6), 2321–2350.
- CANNATA, F., AND M. QUAGLIARIELLO (2005): "The Value of Market Information in Banking Supervision: Evidence from Italy," *Journal of Financial Service Research*, 27(2), 139–162.
- CECCHETTI, S. G., H. GENBERG, AND S. WADHWANI (2002): "Asset Prices in a Flexible Inflation Targeting Framework," NBER working papers, National Bureau of Economic Research.
- COLLIN-DUFRESNE, P., R. S. GOLDSTEIN, AND J. S. MARTIN (2001): "The Determinants of Credit Spread Changes," *The Journal of Finance*, 56(6), 2177–2207.

- COVITZ, D. M., D. HANCOCK, AND M. L. KWAST (2002): *Market Discipline in Banking Reconsidered: The Roles of Deposit Insurance Reform, Funding Manager Decisions and Bond Market Liquidity*. Divisions of Research & Statistics and Monetary Affairs, Federal Reserve Board.
- CROSBIE, P. J., AND J. R. BOHN (2002): “Modeling Default Risk,” *KMV LLC*.
- CURRY, T. J., P. J. ELMER, AND G. S. FISSEL (2007): “Equity market data, bank failures and market efficiency,” *The Journal of Economics and Business*, 59, 536–559.
- DE NICOLÒ, G., AND E. LOUKOIANOVA (2006): “Bank Ownership, Market Structure and Risk,” Working paper, International Monetary Fund.
- DEMIRGÜÇ-KUNT, A., AND E. DETRAGIACHE (2002): “Does Deposit Insurance Increase Banking System Stability? An Empirical Investigation,” *Journal of Monetary Economics*, 49(7), 1373–1406.
- DEMIRGÜÇ-KUNT, A., AND H. HUIZINGA (2004): “Market Discipline and Deposit Insurance,” *Journal of Monetary Economics*, 51(2), 375–399.
- DEMSETZ, R. S., M. R. SAIDENBERG, AND P. E. STRAHAN (1996): “Banks with Something to Lose: The Disciplinary Role of Franchise Value,” *Federal Reserve Bank of New York Economic Policy Review*, 2(2), 1–14.
- DEWATRIPONT, M., AND J. TIROLE (1994): *The Prudential Regulation of Banks*. MIT Press Cambridge, Mass.
- DEYOUNG, R., M. J. FLANNERY, W. W. LANG, AND S. SORESCU (2001): “The Information Content of Bank Exam Ratings and Subordinated Debt Prices,” *Journal of Money, Credit, and Banking*, 33(4), 900–925.
- DIAMOND, D. W., AND P. DYBVIK (1983): “Bank Runs, Deposit Insurance, and Liquidity,” *Journal of Political Economy*, 91(3), 401–419.
- DRIESSEN, J. (2005): “Is Default Event Risk Priced in Corporate Bonds?,” *Review of Financial Studies*, 18(1), 165–195.
- DUFFIE, D., AND K. J. SINGLETON (2003): *Credit Risk: Pricing, Measurement, and Management*. Princeton University Press.
- ELTON, E. J., M. J. GRUBER, D. AGRAWAL, AND C. MANN (2001): “Explaining the Rate Spread on Corporate Bonds,” *Journal of Finance*, 56(1), 247–277.
- EVANOFF, D. D., AND L. D. WALL (2001): “SND Yield Spreads as Bank Risk Measures,” *Journal of Financial Services Research*, 20.

Bibliography

- (2002): “Measures of Riskiness of Banking Organizations: Subordinated Debt Yields, Risk-Based Capital, and Examination Ratings,” *Journal of Banking and Finance*, 26.
- FACCHINETTI, M. (2007): “The Prospects of Market Discipline for Banks. Evidence from the Swiss Bond Market,” mimeo.
- FAMA, E. F., AND K. R. FRENCH (1993): “Common Risk Factors in the Returns on Stocks and Bonds,” *Journal of Financial Economics*, 33(1), 3–56.
- FAN, R., J. G. HAUBRICH, P. RITCHKEN, AND J. B. THOMSON (2003): “Getting the Most Out of a Mandatory Subordinated Debt Requirement,” *Journal of Financial Services Research*, 24(2/3), 149–179.
- FLANNERY, M. J. (1998): “Using Market Information in Prudential Bank Supervision: A Review of the U.S. Empirical Evidence,” *Journal of Money, Credit, and Banking*, 30(3), 273–305.
- (2001): “The Faces of “Market Discipline”,” *Journal of Financial Services Research*, 20(2/3), 107–119.
- FLANNERY, M. J., AND S. M. SORESCU (1996): “Evidence of Bank Market Discipline in Subordinated Debenture Yields: 1983-1991,” *Journal of Finance*, 51(4), 1347–1377.
- GABBI, G., AND A. SIRONI (2005): “Which Factors Affect Corporate Bonds Pricing? Empirical Evidence from Eurobonds Primary Market Spreads,” *The European Journal of Finance*, 11(1), 59–74.
- GLOSTEN, L. R. (1989): “Insider Trading, Liquidity, and the Role of the Monopolist Specialist,” *The Journal of Business*, 62(2), 211–235.
- (1994): “Is the Electronic Open Limit Order Book Inevitable?,” *The Journal of Finance*, 49(4), 1127–1161.
- GONZÁLEZ, F. (2005): “Bank Regulation and Risk-Taking Incentives: An International Comparison of Bnk Risk,” *Journal of Banking and Finance*, 29(5), 1153–1184.
- GORTON, G., AND A. M. SANTOMERO (1990): “Market Discipline and Bank Subordinated Debt: Note,” *Journal of Money, Credit, and Banking*, 22(1), 119–128.
- GOYAL, V. K. (2005): “Market Discipline of Bank Risk: Evidence from Subordinated Debt Contracts,” *Journal of Financial Intermediation*, 14(3), 318–350.
- GREENSPAN, A. (2001): *Harnessing Market Discipline*. The Region, Banking and Policy Issues Magazine, Federal Reserve Bank of Minneapolis.
- GROPP, R., AND J. VESALA (2001): “Deposit Insurance and Moral Hazard: Does the Counterfactual Matter?,” Working Paper No. 47, European Central Bank.

- GROPP, R., J. M. VESALA, AND G. VULPES (2006): “Equity and Bond Market Signals as Leading Indicators of Bank Fragility,” *Journal of Money, Credit, and Banking*, 38(2), 399–428.
- GROSSMAN, S. J. (1976): “On the Efficiency of Competitive Stock Markets Where Trades Have Diverse Information,” *Journal of Finance*, 31(2), 573–85.
- GROSSMAN, S. J., AND J. E. STIGLITZ (1980): “On the Impossibility of Informationally Efficient Markets,” *American Economic Review*, 70(3), 393–408.
- GUNTHER, J. W., M. E. LEVONIAN, AND R. R. MOORE (2001): “Can the Stock Market Tell Bank Supervisors Anything They Don’t Already Know?,” *Federal Reserve Bank of Dallas Economic and Financial Review*, 2nd Quarter, 2–9.
- HANCOCK, D., AND M. L. KWAST (2001): “Using Subordinated Debt to Monitor Bank Holding Companies: Is it Feasible?,” *Journal of Financial Services Research*, 2/3, 147–187.
- HAUNER, D. (2005): “Explaining Efficiency Differences among Large German and Austrian Banks,” *Applied Economics*, 37(9), 969–980.
- HAUSMAN, J. A. (1978): “Specification Tests in Econometrics,” *Econometrica*, 46(6), 1251–1271.
- HELLWIG, M. F. (1980): “On the aggregation of information in competitive markets,” *Journal of Economic Theory*, 22(3), 477–498.
- HERRING, R. J. (2004): “The subordinated debt alternative to Basel II,” *Journal of Financial Stability*, 1(2), 137–155.
- HO, T. S., AND H. R. STOLL (1983): “The Dynamics of Dealer Markets Under Competition,” *The Journal of Finance*, 38(4), 1053–1074.
- HULL, J., AND A. WHITE (1987): “The Pricing of Options on Assets with Stochastic Volatilities,” *The Journal of Finance*, 42(2), 281–300.
- IANNOTTA, G., G. NOCERA, AND A. SIRONI (2007): “Ownership Structure, Risk and Performance in the European Banking Industry,” *Journal of Banking and Finance*, 31(7), 2127–2149.
- IMAI, M. (2007): “The Emergence of Market Monitoring in Japanese Banks: Evidence from the Subordinated Debt Market,” *Journal of Banking and Finance*, 31(5), 1441–1460.
- JONES, E. P., S. P. MASON, AND E. ROSENFELD (1984): “Contingent Claims Analysis of Corporate Capital Structures: An Empirical Investigation,” *The Journal of Finance*, 39(3), 611–625.
- KEALHOFFER, S. (2003): “Quantifying Credit Risk I: Default Prediction,” *Financial Analysts Journal*, 59(1), 30–44.
- KEELEY, M. C. (1990): “Deposit Insurance, Risk, and Market Power in Banking,” *The American Economic Review*, 80(5), 1183–1200.

Bibliography

- KRAINER, J., AND J. A. LOPEZ (2004): "Incorporating Equity Market Information into Supervisory Monitoring Models," *Journal of Money, Credit, and Banking*, 36(6), 1043–67.
- KRISHNAN, C., P. H. RITCHKEN, AND J. B. THOMSON (2006): "On Credit Spread Slopes and Predicting Bank Risk," *Journal of Money, Credit, and Banking*, 38(6), 1545–1574.
- KRUGMAN, P. (1991): "Target Zones and Exchange Rate Dynamics," *Quarterly Journal of Economics*, 106(3), 669–682.
- KYLE, A. S. (1985): "Continuous Auctions and Insider Trading," *Econometrica*, 53(6), 1315–35.
- (1989): "Informed Speculation with Imperfect Competition," *The Review of Economic Studies*, 56(3), 317–355.
- LA PORTA, R., F. LOPEZ-DE-SILANES, AND A. SHLEIFER (2002): "Government Ownership of Banks," *The Journal of Finance*, 57(1), 265–301.
- LEHAR, A., D. SEPPI, AND G. STROBL (2006): "Using Price Information as an Instrument of Market Discipline in Regulating Bank Risk," Working papers, University of Calgary.
- LEHMANN, H., AND M. MANZ (2006): "The Exposure of Swiss Banks to Macroeconomic Shocks - An Empirical Investigation," SNB Working Papers 2006-4, Swiss National Bank.
- LONGHOFFER, S. D., AND C. T. CARLSTROM (1995): "Absolute Priority Rule Violations in Bankruptcy," *Federal Reserve Bank of Cleveland Economic Review*, 31(4), 21–30.
- LONGSTAFF, F. A., AND E. S. SCHWARTZ (1995): "A Simple Approach to Valuing Risky Fixed and Floating Rate Debt," *The Journal of Finance*, 50(3), 789–819.
- MARCUS, A. J. (1984): "Deregulation and Bank Financial Policy," *Journal of Banking and Finance*, 8(4), 557–565.
- MARSHALL, D. A., AND E. S. PRESCOTT (2001): "Bank Capital Regulation with and Without State-Contingent Penalties," *Carnegie-Rochester Conference Series on Public Policy*, 54(1), 139–184.
- MERTON, R. C. (1974): "On the Pricing of Corporate Debt: The Risk Structure of Interest Rates," *Journal of Finance*, 29(2), 449–70.
- (1977): "An Analytic Derivation of the Cost of Deposit Insurance and Loan Guarantees," *Journal of Banking and Finance*, 1(1), 3–11.
- MILGROM, P., AND N. STOKEY (1982): "Information, trade and common knowledge," *Journal of Economic Theory*, 26(1), 17–27.
- MORGAN, D. P., AND K. J. STIROH (2001): "Market Discipline of Banks: The Asset Test," *Journal of Financial Services Research*, 20(2), 195–208.

- MORRIS, S., AND H. S. SHIN (2004): "Coordination Risk and the Price of Debt," *European Economic Review*, 48(1), 133–153.
- NIER, E., AND U. BAUMANN (2006): "Market Discipline, Disclosure and Moral Hazard in Banking," *Journal of Financial Intermediation*, 15(3), 332–361.
- PARK, S., AND S. PERISTIANI (2007): "Are bank shareholders enemies of regulators or a potential source of market discipline?," *Journal of Banking and Finance*, 31(8), 2493–2515.
- PERSSON, M. (2002): "Stock Market Indicators of Bank Fragility - Evidence from the Nordic Banking Crisis," mimeo, Sveriges Riksbank.
- POP, A. (2006): "Market Discipline in International Banking Regulation: Keeping the Playing Field Level," *Journal of Financial Stability*, 2(3), 286–310.
- POSEN, A. S. (2006): "Why Central Banks Should Not Burst Bubbles," *International Finance*, 9(1), 109–124.
- RIME, B. (2005): "Do Too Big to Fail Expectations Boost Large Banks Issuer Ratings?," mimeo, Swiss National Bank.
- RIME, B., AND K. J. STIROH (2003): "The performance of universal banks: Evidence from Switzerland," *Journal of Banking and Finance*, 27(11), 2121–2150.
- ROUBINI, N. (2006): "Commentary on "Why Central Banks Should Burst Bubbles"," *International Finance*, 1, 87–107.
- SAUNDERS, A., AND B. WILSON (2001): "An Analysis of Bank Charter Value and Its Risk-Constraining Incentives," *Journal of Financial Services Research*, 19(2/3), 185–195.
- SHIN, H. S. (2002): "Comments on "How good is the market at assessing bank fragility? A horse race between different indicators"," *Journal of Banking and Finance*, 26(5), 1029–1031.
- SIRONI, A. (2001): "An analysis of European banks SND issues and its implications for the design of a mandatory subordinated debt policy," *Journal of Financial Services Research*, 20(2/3), 233–266.
- (2002): "Strengthening banks' market discipline and leveling the playing field: Are the two compatible?," *Journal of Banking and Finance*, 26(5), 1065–1091.
- (2003): "Testing for Market Discipline in the European Banking Industry: Evidence from Subordinated Debt Issues," *Journal of Money, Credit, and Banking*, 35(3), 443–72.
- STEVEY, R. (2007): "Bank Size, Credit and the Sources of Bank Market Risk," BIS Working Papers 238, Bank for International Settlements.

Bibliography

- STIGLITZ, J. E., AND A. WEISS (1981): "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, 71(3), 393–410.
- STOLL, H. R. (1978): "The Supply of Dealer Services in Securities Markets," *The Journal of Finance*, 33(4), 1133–1151.
- SWIDLER, S., AND J. A. WILCOX (2002): "Information about bank risk in options prices," *Journal of Banking and Finance*, 26(5), 1033–1057.
- SY, A. N. (2002): "Emerging Market Bond Spreads and Sovereign Credit Ratings: Reconciling Market Views with Economic Fundamentals," *Emerging Markets Review*, 3(4), 380–408.
- VAN LANDSCHOOT, A. (2004): "The Determinants of Credit Spreads," *Financial Stability Review*, pp. 135–155.
- VERRECCHIA, R. E. (1982): "Information Acquisition in a Noisy Rational Expectations Economy," *Econometrica*, 50(6), 1415–1430.
- VON HAYEK, F. A. (1945): "The Use of Knowledge in Society," *American Economic Review*, 35(4), 519–530.
- WOLFERS, J., AND E. ZITSEWITZ (2004): "Prediction Markets," *Journal of Economic Perspectives*, 18(2), 107–126.
- WOOLDRIDGE, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA.